UNIVERSIDADE ESTADUAL DE
CAMPINAS

Instituto de Matemática, Estatística e
Computação Científica

GUSTAVO DO NASCIMENTO COSTA

Equações Diferenciais Estocásticas Backward: Uma Aplicação em Finanças

Campinas

2018

Gustavo do Nascimento Costa

Equações Diferenciais Estocásticas Backward: Uma Aplicação em Finanças

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Matemática.

Orientador: Pedro José Catuogno

Este exemplar corresponde à versão final da Dissertação defendida pelo aluno Gustavo do Nascimento Costa e orientada pelo Prof. Dr. Pedro José Catuogno.

Campinas

2018

Agência(s) de fomento e nº(s) de processo(s): CNPq, 131874/2016-0

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

C823e
Costa, Gustavo do Nascimento, 1993-
Equações diferenciais estocásticas backward : uma aplicação em finanças / Gustavo do Nascimento Costa. – Campinas, SP : [s.n.], 2018.

Orientador: Pedro José Catuogno.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Itô, Cálculo de. 2. Equações diferenciais estocásticas backward. 3. Feynman-Kac, Representação de. 4. Black-Scholes, Modelo de. I. Catuogno, Pedro José, 1959-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Backward stochastic differential equations : a financial application
Palavras-chave em inglês:
Itô calculus
Backward stochastic differential equations
Feynman-Kac representation
Black-Scholes model
Área de concentração: Matemática
Titulação: Mestre em Matemática
Banca examinadora:
Pedro José Catuogno [Orientador]
Roberto Andreani
Dorival Leão Pinto Júnior
Data de defesa: 15-03-2018
Programa de Pós-Graduação: Matemática

Dissertação de Mestrado defendida em 15 de março de 2018 e aprovada pela banca examinadora composta pelos Profs. Drs.

Prof(a). Dr(a). PEDRO JOSE CATUOGNO

Prof(a). Dr(a). ROBERTO ANDREANI

Prof(a). Dr(a). DORIVAL LEÃO PINTO JÚNIOR

As respectivas assinaturas dos membros encontram-se na Ata de defesa

À dona Zilda,

Agradecimentos

Primeiramente gostaria de agradecer ao CNPq e ao SAE, por terem me proporcionado condições concretas e dignas de existência para realizar esse estudo. Que sobrevivam e floresçam nessas épocas tão duras.

Agradeço à minha mãe, Maria, por ter moldado meu caráter em meio a tantas adversidades.

Agradeço especialmente e com todas minhas forças ao meu querido amigo Leonardo Francisco Cavenaghi. Não só por ter sido verdadeiro mentor no meu trabalho, me ajudando com toda uma infinidade de burocracias, uma pletora de dificuldades com Latex e uma pequena grande coisa em geometria; eu definitivamente não conheço melhor pessoa a se correr com esses apuros kafkinianos e me pergunto até hoje quem lhe ensinou a lidar com eles. Mas também, e principalmente, por ter estado comigo nos momentos mais obscuros dessa caminhada. Sabemos o que ela foi, e definitivamente sem seu apoio eu não teria terminado. Vivemos as mesmas coisas juntos, e você me mostrou como superá-las. Iluminou-me o caminho. Em verdade, sinto que nunca conseguirei expressar como fui e sou grato a você, à nossa amizade, e a ter me salvado nesse período.

Agradeço ao Allan, cuja companhia, infindáveis conversas, caminhadas e mapas me trouxeram serenidade de espírito. Nunca conheci alguém mais sensível e decente. Realmente nossa amizade trouxe muita luz.

Agradeço à Bia, por todo seu companheirismo, carinho e atenção. Muito obrigado por enxergar beleza nas coisas que me assustam. E também, sem sua mesa eu não teria escrito esse trabalho.

Agradeço ao Ângelo, que me ouviu e me entendeu em momentos complicados. Sempre disse coisas muito valiosas e de calmaria. Atribuo à sua influência minha maneira de ler e argumentar.

Agradeço ao Rafa, que, paradoxalmente, sempre me fez ater aos prazos. E também por ser uma pessoa extremamente humana.

Agradeço ao João pelos muitos anos de amizade e por ser a única pessoa a entender e concordar com minhas visões mais peculiares.

Agradeço ao IMECC e à Unicamp, pela boa estrutura de pesquisa. Mas mais especificamente, agradeço muito à secretaria de pós do IMECC. Sempre muito atenciosos e competentes, realmente dispostos a fazer a roda girar.

Por fim, agradeço ao Pedro, por ter aceitado me orientar e ter me dado um bom tema, e à banca, professores Dorival e Roberto, por terem aceitado ler e avaliar meu trabalho.

"For, after all, how do we know that two and two make four? Or that the force of gravity works? Or that the past is unchangeable? If both the past and the external world exist only in the mind, and if the mind itself is controllable – what then?"

(George Orwell, 1984)

Resumo

Neste trabalho, introduzimos a integral de Itô para processos progressivamente mensuráveis como uma generalização da transformada martingale e construímos o cálculo relacionado. Com essa nova teoria de integração, estudamos um tipo especial de equação diferencial, as chamadas BSDEs (Equações Diferenciais Estocásticas "Backward"), cujas soluções, obtidas através do teorema de representação de martingales, são pares de processos adaptados. Estudamos o teorema padrão sobre existência e unicidade dessas soluções, algumas de suas propriedades e encontramos soluções para dois tipos específicos de BSDEs. Por fim, mostramos (i) como esse tipo de equação é usada dentro do contexto de EDPs a fim de se generalizar a representação de Feynman-Kac, permitindo-nos representar estocasticamente EDPs semilineares; e (ii) como as BSDEs aparecem naturalmente na modelagem de precificação de derivativos. Em especial, usamo-las para encontrar o valor inicial de uma opção de compra europeia de um ativo de risco no mercado, sem nos preocuparmos com sua completude, segundo o modelo de Black-Scholes.

Palavras-chave: Cálculo de Itô. BSDEs. Representação de Feynman-Kac. Modelo de Black-Scholes

Abstract

In this work, we introduce the Itô integral for progressively measurable processes as a generalizaton of the martingale transform, and we build the related calculus. With this new theory of integration, we study a special type of differential equation, the so called BSDEs (Backward Stochastic Differential Equations), whose solutions, obtained through the martingale representation theorem, are pairs of adapted processeses. We study the standard theorem about existence and unicity of such solutions, some of its properties, and we find solutions for two specific kinds of BSDEs. Lastly, we show (i) how this kind of equation is used within the PDE context in order to generalize de Feynman-Kac representation, allowing us to represent stochasticly semilinear PDEs; and (ii) how the BSDEs appear naturally in the modelling of derivative precification. In particular, we use them to find the initial price of a european call option of risk assets in a market, without concerning on whether the market is complete or not, according to the Black-Scholes model.

Keywords: Itô Calculus. BSDEs. Feynman-Kac Representation. Black-Scholes Model.

Sumário

Introdução

0.1 Informação e σ-Álgebra

Esta dissertação de mestrado tem como meta central introduzir as BSDEs. Claro que introduzir um conceito novo em matemática é feito em etapas. Constrói-se as bases necessárias, motiva-se o estudo do novo conceito, define-se com precisão este novo conceito, estuda-se os principais resultados e tenta-se mostrar exemplos e aplicações. Esta é basicamente a linha seguida no texto. A motivação dada ao estudo das BSDEs nasce aqui com o desejo de se generalizar a representação de Feynman-Kac para EDPs semilineares. Vemos que, ao tentar fazer isso, somos levados a um tipo de equação diferencial cuja solução é um **par** de processos **adaptados**. Acontece que, bem embora tenhamos partido de tal motivação, não é a generalização desta representação nosso principal exemplo. As BSDEs são amplamente utilizadas em temas financeiros, seja na precificação de derivativos, como abordado na seção 2.7, seja no tratamento de questões financeiras como problemas de controle estocástico (aliás, enxergar a solução de um problema assim via BSDEs constitui aplicação amplamente difundida deste tipo de equação). Ambas aplicações são abordadas de forma rigorosa em (PHAM, 2009) e (KAROUI; PENG; QUENEZ, 1997). Pode-se dizer então que entender este último artigo foi um dos objetivos deste estudo de mestrado (em verdade, era a meta inicial; contudo, os objetivos foram se alternando conforme conceitos novos foram aparecendo e necessitando esclarecimento). E é justamente aqui que entramos com algumas considerações sobre aplicação de análise estocástica em finanças. Enquanto os trabalhos mais recentes se valem de uma matemática sofisticada para abordar problemas em finanças, algumas questões, aparentemente ingênuas, nasceram bem no início do estudo da interface matemática pura/finanças. Se, por um lado, temos resultados abstratos e rigorosos amplamente utilizados para resolver problemas que pareciam complicados demais, e estudar tal matemática confere assunto suficiente para algo muito mais longo que essa dissertação, por outro foi uma pergunta singela, emergida nas premissas da modelagem, que ocupou grande parte do estudo. E, embora o texto passe longe dessa questão, creio que uma introdução a comportaria bem. E a questão é: por que os processos que modelam um portfólio tem de ser adaptados?

Um processo $X : [0,T] \times \Omega \to \mathrm{R}^n$ é dito **adaptado** a uma filtração $\{\mathcal{F}_t\}$ quando X_t for $\mathcal{F}_t$ mensurável para cada $t \in [0,T]$. Diz-se extensivamente que essa definição modela um processo X cujos valores, em cada instante t, são condizentes com a **quantidade de informação** contida na σ-álgebra $\mathcal{F}_t$. Assim, necessariamente nascem duas novas questões: 1. Como possivelmente estruturas abstratas como σ-álgebras modelam o conceito de informação; 2. Como uma variável aleatória $\mathcal{F}$-mensurável "condiz" com a quantidade de informação contida em $\mathcal{F}$.

Essa duas perguntas são respondidas de forma extremamente satisfatória em (HERVéS-BELOSO; MONTEIRO, 2012), e seguimos com uma adaptação de suas ideias. Considere um espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$. Pela ótica da teoria da medida, ele é tratado como um espaço de medida de medida unitária, de forma que ao se considerar variáveis aleatórias $f : \Omega \to \mathrm{R}$, simplesmente chamadas de funções mensuráveis, dá-se o foco usual ao domínio Ω. Acontece que, numa abordagem probabilística, usualmente há um interesse quase que exclusivo nas distribuições de probabilidade das variáveis aleatórias, seja na medida induzida em R por f, seja na sua distribuição acumulada, de forma que muitas vezes estuda-se essas variáveis sem sequer mencionar o espaço Ω em que são definidas (o que não é necessariamente uma adversidade, dado o tipo de problemas que se aborda em probabilidade). Curiosamente, o grande mecanismo para se entender como a σ-álgebra $\mathcal{F}$ contém informação é voltar ao espaço Ω e enxergá-lo como um **espaço de estados**, de forma que, agora, a v.a. f atribui valores a cada **estado do mundo** $\omega \in \Omega$, e perceber que nos importamos com esses estados. A princípio, vamos supor que Ω é finito. Seja $\mathcal{P}$ o conjunto de partições do espaço Ω e $\mathcal{S}$ o conjunto de σ-álgebras de Ω. Temos o seguinte teorema.

Teorema 0.1.1. *Há uma bijeção $\phi : \mathcal{P} \to \mathcal{S}$ tal que $P_2 \in \mathcal{P}$ é mais refinada que $P_1 \in \mathcal{P}$ se e somente se $\phi(P_1) \subset \phi(P_2)$.*

Demonstração. Seja $\phi : \mathcal{P} \to \mathcal{S}$ dada por $P \to \sigma(P)$. Da definição, vê-se claramente que vale a relação entre refinamento de partição e continência de σ-álgebras enunciada. Vamos então definir uma função $\psi : \mathcal{S} \to \mathcal{P}$ que mostraremos ser a inversa de ϕ. Para $\mathcal{F} \in \mathcal{S}$ e $x \in \Omega$, considere o conjunto

$$A_x^{\mathcal{F}} := \bigcap_{\substack{E \in \mathcal{F} \\ x \in E}} E.$$

Notemos que a relação $x \sim_{\mathcal{F}} y :\Leftrightarrow y \in A_x^{\mathcal{F}}$ é uma relação de equivalência em Ω. Essa relação então define uma partição $P_{\mathcal{F}}$ em Ω. Definimos assim $\psi(\mathcal{F}) = P_{\mathcal{F}}$. Mostremos que ψ e ϕ são inversas:

1. Seja $P = \{P_1, \ldots, P_n\} \in \mathcal{P}$. Então $\sigma(P) = \{\bigcup_{i \in I} P_i\}_{I \subseteq \{1,\ldots,n\}}$. Dessa forma, fica claro que os blocos da partição induzida por $\sigma(P)$ são precisamente os blocos de P. Isso mostra que $\psi \circ \phi = \mathrm{Id}_{\mathcal{P}}$.

2. Seja agora $\mathcal{F} \in \mathcal{S}$. Considere a sigma álgebra $\mathcal{G}$ gerada pela partição $P_{\mathcal{F}}$. Como cada elemento da partição $P_{\mathcal{F}}$ está claramente em $\mathcal{F}$ (por ser uma intersecção finita), temos $\mathcal{G} \subseteq \mathcal{F}$, trivialmente. Seja $E \in \mathcal{F}$. É fácil ver que $E = \cup_{x \in E} A_x^{\mathcal{F}}$. Isso nos diz que E é uma união finita de elementos da partição $P_{\mathcal{F}}$, ou seja, $E \in \sigma(P_{\mathcal{F}})$. Assim, $\mathcal{F} \subseteq \mathcal{G} \Rightarrow \mathcal{F} = \mathcal{G}$. Isso nos diz que $\phi \circ \psi = \mathrm{Id}_{\mathcal{S}}$. Nosso teorema segue. $\square$

Dada uma partição P, chamamos costumeiramente de **σ-álgebra gerada por** P a σ-álgebra $\sigma(P)$. Chamaremos então de **partição gerada por** σ a $\phi^{-1}(\sigma)$, onde ϕ é a função do teorema 0.1.1. Agora, temos também a seguinte definição:

Definição 0.1.2. *Um sinal para Ω é uma função $f : \Omega \to \mathrm{R}$.*

Notemos que sinais para Ω geram partições em Ω, onde cada $P_i = f^{-1}(x), x \in \mathrm{R}$. Suponha então que o estado do mundo $\omega \in \Omega$ é algo desconhecido, e que tentamos descobri-lo através de sinais (ou medições) f. Quanto mais refinada a partição gerada por f, mais certeza sobre qual o real estado do mundo obteremos. Isso nos cria a ideia de *informação gerada por uma variável aleatória*. São modelos de informação obtida através da **observação** de determinada medição, ou sinal, ou v.a., que nos aproximam mais do conhecimento exato do real estado do mundo considerado, via partição obtida. Assim, dizemos que uma v.a. **gera** informação $\mathcal{F}$ (e usamos uma letra de σ-álgebra) quando consideramos primeiramente a partição por ela gerada, algo feito através da definição da nossa v.a. (ou sinal), e posteriormente uma simulação dessa v.a. (uma medição) nos apontará em que possíveis estados de mundo realmente estamos. Vamos entender agora o que significa uma v.a. ser **condizente** com a quantidade de informação obtida numa σ-álgebra. Para isso, considere o seguinte resultado:

Proposição 0.1.3. *Uma função $g : \Omega \to \mathrm{R}$ é $\mathcal{F}$-mensurável se, e só se, g for constante na partição gerada por $\mathcal{F}$.*

E é precisamente isso que nos dá a ideia de função g adaptada à quantidade de informação contida em $\mathcal{F}$. Se g representar, por exemplo, uma tomada de decisão que é determinada unicamente por cada estado de mundo ω, ter-se uma σ-álgebra $\mathcal{F}$ (gerada por um sinal f, por exemplo) pode ser entendido como ter uma forma de se descobrir quais estados de mundo são possíveis (através do nosso sinal ou medição f, neste caso), de forma que g é condizente com a quantidade de informação dada por $\mathcal{F} = \sigma(f)$ se e só for $\mathcal{F}$-mensurável, ou seja, se e só se, para cada conjunto de estados de mundo possíveis, obtidos através de f, g for bem definido e constante, o que reflete uma decisão que contempla, simultaneamente, todos os estados contidos no conjunto da partição. Seria algo como: antes de se tomar uma decisão g, realiza-se uma medição f a fim de se descobrir, mais ou menos, em qual estado o mundo se encontra. Sabe-se com que grau de acurácia f é capaz de fazer isso, ou seja, conhece-se a partição em Ω gerada por f. Decide-se então um valor fixo para g em cada elemento da partição, ou seja, pensa-se qual seria a melhor decisão em cada um dos cenários possíveis de conjunto de estados de mundo a ser considerados. Ao se realizar a medição f termina-se conhecendo em quais estados de mundo de fato pode-se estar, e, portanto, sabe-se qual decisão tomar. Essa relação entre informação fornecida por uma σ-álgebra e informação condizente com uma σ-álgebra é sumarizada pelo teorema seguinte.

Teorema 0.1.4. *Seja $X : \Omega \to \mathrm{R}$. Então, X é $\sigma(f)$-mensurável se, e só se, $X = \phi \circ f$, onde $\phi : \mathrm{R} \to \mathrm{R}$ é Borel-mensurável.*

Demonstração. Ver (BOULEAU, 1986, pág 101-102). □

Assim, uma decisão é condizente com a informação gerada por um sinal se, e só se, for uma função do sinal, o que parece bem óbvio, e encerramos assim a nossa leitura no caso em que Ω é finito.

A questão agora é como a relação entre σ-álgebra e informação se traduz para o contexto onde Ω não é finito. Notando que essa relação nasce da dualidade partição-σ-álgebra, basicamente precisamos traduzir o teorema 0.1.1 para o caso de Ω infinito. E esse é precisamente o conteúdo de (HERVéS-BELOSO; MONTEIRO, 2012, Seção 5.1), que apresenta a **teoria de Blackwell**, generalizadora dessas ideias para espaços infinitos. Em particular, nesta dissertação nós trabalhamos com a informação advinda de um movimento Browniano, definido num espaço possivelmente grande demais. Acontece que, de acordo com A.2, podemos definir o movimento Browniano (ou qualquer outro processo contínuo), num espaço de funções contínuas $\Omega := C([0, T]; \mathrm{R})$, que é suficientemente "bem comportado" a ponto de se enquadrar na teoria descrita, de forma que nossa interpretação intuitiva sobre informação/σ-álgebra mantém-se intacta neste novo cenário.

0.2 Martingales

O termo martingale tem origem num conjunto de estratégias de aposta comuns na França no século 18. Num jogo que consistia no lançamento de uma moeda em repetidos turnos, procurava-se uma forma de alocar recursos anteriormente a cada turno de forma a conseguir uma vantagem no jogo em seu desenlace. Acontece que os jogos justos não permitem que o conhecimento do comportamento passado da moeda conceda a alguém o poder de inferir com mais certeza sobre seu proceder futuro. E é justamente por essa observação que estratégias de aposta, por mais bem elaboradas, não permitiam um ganho indubitável em jogos justos. Isso vem a ser formalizado com teorema do tempo de parada de Doob, A.3.1, que versa sobre a impossibilidade de parar o jogo de maneira esperta o suficiente para se obter certamente algum lucro, e com o a proposição (1.1.1), que fala sobre as estratégias de aposta. O estudo de estratégias de investimento possui um paralelo com estratégias de apostas, onde considera-se o investidor um apostador, como na seção 1.1. Com isso, de posse da discussão sobre como σ-álgebras modelam informação, nasceu, neste estudo, a dúvida sobre a relação entre a noção de um jogo em que o *conhecimento* do passado não permite predições sobre um comportamento futuro (onde conhecimento agora pode ser tratado via σ-álgebras) e a definição matemática de um martingale, dada por:

Definição 0.2.1. *Um processo estocástico* $(X_n)_{n\in\mathrm{N}}$ *definido no espaço de probabilidade filtrado* $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathrm{N}}, \mathrm{P})$ *e cujas todas v.a.'s são integráveis, é dito um martingale se*

$$\mathrm{E}[X_m|\mathcal{F}_n] = X_n, \qquad \forall m \geqslant n \in \mathrm{N}. \tag{1}$$

Se X_t *for um processo a tempos contínuos em* $[0, T]$*, num espaço com filtração também a tempos contínuos, a definição se mantém, trocando* $m, n \in \mathrm{N}$ *em 1 acima por* $s \geqslant t \in [0, T]$*, respectivamente.*

A pergunta então é: o que é um martingale? A breve noção intuitiva que daremos é fortemente baseada no artigo homônimo (DOOB, 1971). Não há como responder essa questão sem entender exatamente o conceito de esperança condicional. Lembremos que ela é definida por

Definição 0.2.2. *Dado um espaço de probabilidade* $(\Omega, \mathcal{F}, \mathrm{P})$*, uma* σ*-álgebra* $\mathcal{G} \subseteq \mathcal{F}$*, e uma v.a.* $\mathcal{F}$*-mensurável e integrável* X*, a esperança condicional de* X *em relação a* $\mathcal{G}$*, denotada por* $\mathrm{E}[X|\mathcal{G}]$*, é a única variável aleatória* $\mathcal{G}$*-mensurável tal que*

$$\mathrm{E}[X; A] = \mathrm{E}[\mathrm{E}[X|\mathcal{G}]; A],$$

para todo $A \in \mathcal{G}$*.*

A existência e unicidade qtp de uma esperança condicional (para v.a.'s integráveis) são garantidas pelo teorema de Radon–Nikodym. É claro que se a v.a. X for quadrado integrável, isto é, se $X \in L^2(\Omega, \mathcal{F}, \mathrm{P})$, a esperança condicional é caracterizada por

$$\mathrm{E}[X|\mathcal{G}] = \mathrm{proj}^{\perp}_{L^2(\Omega,\mathcal{G},\mathrm{P})} X. \tag{2}$$

Assim, suponha então, para facilitar a visualização, que Ω, nosso conjunto de estados, seja finito. Suponha também que temos duas σ-álgebras $\mathcal{F}$ e $\mathcal{G}$ tais que $\mathcal{F} \supset \mathcal{G}$. Isso significa que a partição $P(\mathcal{F})$ gerada por $\mathcal{F}$ é mais refinada que a partição $P(\mathcal{G})$ gerada por $\mathcal{G}$. Seja X uma v.a. $\mathcal{F}$-mensurável integrável. Então X é constante nos elementos de $P(\mathcal{F})$. Vamos calcular $E[X|\mathcal{G}]$. Seja $A \in P(\mathcal{G})$. Então $A = A_1 \sqcup \ldots \sqcup A_n$, com cada $A_i \in P(\mathcal{F})$, e escrevemos C_{A_i} para o valor de X em cada A_i. Note que:

- $\mathrm{E}[X|\mathcal{G}]$ é $\mathcal{G}$-mensurável, portanto constante $(= C_A)$ em A;
- $\mathrm{E}[X1_A] = \mathrm{E}[\mathrm{E}[X|\mathcal{G}]1_A]$.

Com isso, temos

$$C_A\mathrm{P}(A) = \mathrm{E}[X1_A] \Rightarrow C_A = \frac{\mathrm{E}[X1_A]}{\mathrm{P}(A)} = \frac{C_{A_1}\mathrm{P}(A_1)}{\mathrm{P}(A)} + \ldots + \frac{C_{A_n}\mathrm{P}(A_n)}{\mathrm{P}(A)} = \mathrm{E}[X|A]. \tag{3}$$

Então, a esperança condicional de X em relação a $\mathcal{G}$ é uma aproximação (a melhor em L^2, por 2, para v.a.'s quadrado integráveis) para a variável X, se enquadrando na restrição de ser constante em cada elemento da partição de $\mathcal{G}$. Ou seja, para saber o valor de $\mathrm{E}[X|\mathcal{G}]$ no elemento A da partição de $\mathcal{G}$, calculamos a esperança de X **dado que** A aconteceu. Podemos agora entender o que é um martingale e porque modela jogos honestos. Considere um espaço filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathrm{P})$ e um processo adaptada à filtração, V_n. A filtração nos diz que conhecemos de maneira cada vez mais precisa o real estado do mundo e que, sabendo-se em que elemento da partição de $\mathcal{F}_n$ o mundo se encontra, sabemos o valor da grandeza V_n. Dizer que $\mathrm{E}[V_n|\mathcal{F}_{n-1}] = V_{n-1}$ significa dizer que, de posse do conjunto A em que o estado do mundo se encontra no instante $n-1$, sempre de acordo com a corpulência das informações obtidas em tal instante, e portanto conhecendo exatamente o valor de V_{n-1}, esperamos de V venha a valer no instante n não outra coisa que o valor de V_{n-1}. **Não que V_n deva assumir o valor de V_{n-1} dentro do conjunto** A. Na verdade, V_n pode ter, em A, tanta diversificação de valores quanta for a quantidade de elementos da subpartição de A gerada por $\mathcal{F}_n$. O que acontece é que, **na média**, V_n assume em A o valor (constante) de V_{n-1}. É como se dentro de um conjunto de estados de mundo A, as grandezas V_n vão se especializando mais e mais a cada instante, sem perder seu valor médio em A. Mostremos como isso modela precisamente um jogo justo. Para tanto, vamos considerar o exemplo de um jogador que ganha ou perde 1 real conforme o resultado de uma moeda honesta arremessada é cara ou coroa. Se $(X_n)_{n=1}^{\infty}$ for um conjunto de v.a.'s independentes e identicamente distribuídas por $\mathrm{P}(X_i = 1) = \mathrm{P}(X_i = -1) = 1/2$, então estamos interessados nas strings de valores 1 e -1, representando os resultados dos arremessos de moeda. Nosso conjunto de estados Ω é dado portanto por elementos do tipo $(x_n)_{n=1}^{\infty}$, $x_n \in \{1, -1\}$, e $V_n = X_1 + \ldots + X_n$ modela a riqueza do jogador em cada instante n. Note que a cada instante n, após o arremesso da moeda, temos uma nova σ-álgebra $\mathcal{F}_n$ em Ω, gerada por $X_1, \ldots, X_n$. Aí nos perguntamos: dado que a n rodada do jogo se foi, qual a expectativa em relação à riqueza do jogador na rodada $n+1$? Ou seja, sabendo exatamente o que aconteceu com a moeda até os instante n, quanto o jogador espera ganhar? Vamos calcular isso de duas maneiras:

1. Queremos saber $\mathrm{E}[V_{n+1}|X_1, \ldots, X_n]$. Tomamos então o estado de mundo A em que o jogo se encontra no instante n, ou seja, os valores já obtidos por $X_1, \ldots, X_n$ (note que há muitos elementos em A, representando todos os possíveis valores da moeda nos instantes futuros a n, respeitando o passado até n já ocorrido). A será um conjunto do tipo $\{X_1 = x_1, \ldots, X_n = x_n\}$, com $x_i = 1$ ou -1. Então, por (3), temos

$$
\begin{aligned}
&\mathrm{E}[V_{n+1}|X_1, \ldots, X_n] = V(X_1 = x_1, \ldots, X_n = x_n, X_{n+1} = 1)\mathrm{P}(X_{n+1} = 1) + \\
&V(X_1 = x_1, \ldots, X_n = x_n, X_{n+1} = -1)\mathrm{P}(X_{n+1} = -1) \\
&= \frac{1}{2}(x_1 + \ldots + x_n + 1) + \frac{1}{2}(x_1 + \ldots + x_n - 1) = V_n,
\end{aligned}
$$

em A. Ou seja, a grandeza é um martingale.

2. Por propriedades da esperança condicional, $\mathrm{E}[V_{n+1}|\mathcal{F}_n] = X_1 + \ldots + X_n + \mathrm{E}[X_{n+1}|\mathcal{F}_n] = X_1 + \ldots + X_n + \mathrm{E}[X_{n+1}] = X_1 + \ldots + X_n = V_n$, ou seja, V_n é mesmo um martingale.

O interessante é que, justamente através do maquinário matemático de martingales, podemos mostrar que se o jogador bolar algum tipo de estratégia para jogar em cada instante n que envolva apenas seu conhecimento do estado de mundo em $n-1$, e usamos aqui a compreensão σ-álgebra/informação, sua riqueza final V_n continuará sendo um martingale (veja a proposição 1.1.1). Isso traduz precisamente, em termos matemáticos, a noção intuitiva de que não se pode bolar uma estratégia esperta de apostas que permitam o jogador esperar ganhar dinheiro baseado apenas em informação passada que não interfere no comportamento futuro do jogo.

0.3 Movimento Browniano

É justamente do estudo de jogos justos como o arremesso de uma moeda, cuja riqueza associada V_n (na notação da seção anterior) é usualmente chamada de **passeio aleatório em dimensão** 1, que nasce a resposta da última questão incipiente sobre a modelagem feita por Black-Scholes, e largamente utilizada num contexto financeiro, sobre o preço de um ativo de risco no mercado, digamos, uma ação. De antemão, anunciamos que a modelagem envolve um movimento Browniano, dizendo que se B_t representa um ativo de risco no mercado, então seu preço é dado por

$$B_t = B_0 \exp{(\mu - \frac{\sigma^2}{2})t + \sigma dW_t},$$

onde W_t é um **movimento Browniano unidimensional**. Na verdade, é bem mais comum ver a seguinte expressão na modelagem:

$$dB_t = B_t(\mu dt + \sigma dW_t).$$

A questão portanto é: por que usa-se o movimento Browniano para modelar o preço de B_t? E a resposta: basicamente porque enxergamos a parte aleatória dos ganhos de uma ação como um passeio aleatório “continuizado”, como se jogássemos uma moeda a todo instante para decidir se esse preço deveria subir ou descer, não mais uma unidade, já que temos jogos novos a todos instantes, mas um valor de alguma forma proporcional. Falamos mais sobre isso na seção 1.1. Precisamos então entender como o movimento Browniano é enxergado como uma “continuização” de um passeio aleatório. Vamos seguir quase literalmente a abordagem dada em (KUO, 2005, Seção 1.2), que apesar de claramente prezar pela intuição, segue uma sequência com razoável rigor. Consideramos o intervalo $[0, T]$ dividido em M subintervalos de tamanhos idênticos $\delta = T/M$. Consideramos também uma sequência de variáveis aleatórias X_i independentes e igualmente distribuídas por $\mathrm{P}(X_i = h) = \mathrm{P}(X_i = -h) = 1/2$ e chamamos h de *tamanho do passo* do passeio aleatório definido por X_i. Queremos construir um processo Y em $[0, T]$ que simule diversos arremessos de moeda, e faremos isso considerando o limite de um passeio aleatório com cada vez mais arremessos de passos cada vez menores. Começamos então dividindo $[0, T]$ em M subintervalos de tamanho idênticos igual a $\delta = T/M$. Então definimos uma sequência de processos $Y^{\delta,h}$ em $[0, T]$ por

$$Y^{\delta,h}(n\delta) = \sum_{i-1}^{n} X_i,$$

ou seja, somas de n moedas de passo h jogadas nos instantes $n\delta$ $(n = 0, 1, \ldots, M)$, e linearizamos cada trajetória de $Y^{\delta,h}$ no resto do intervalo:

$$Y^{\delta,h}(t) = \frac{(n+1)\delta - t}{\delta} Y^{\delta,h}(n\delta) + \frac{t - n\delta}{\delta} Y^{\delta,h}((n+1)\delta),$$

para $n\delta < t < (n+1)\delta$. E agora nossa ideia é justamente estudar o limite de $Y^{\delta,h}$ quando $(\delta, h) \to (0, 0)$, ou seja, quando o tamanho do passo e tempo de jogada tendem a 0. Estudemos a função característica de cada $Y^{\delta,h}(t)$, já que estas determinam unicamente varáveis aleatórias. Avaliemos estas funções nos instantes $t = n\delta$:

$$\mathrm{E}[\exp(i\lambda Y^{\delta,h})(t)] = \prod_{j=1}^{n} \mathrm{E}[\exp(i\lambda X_j))] =$$

$$(\mathrm{E}[\exp(i\lambda X_1)^n]) = (\frac{1}{2}e^{i\lambda h} + \frac{1}{2}e^{-i\lambda h})^n = (\cos\lambda h)^n = (\cos\lambda h)^{t/\delta}.$$

onde a primeira igualdade vem da independência das X_j, a segunda do fato de as X_j terem a mesma distribuição e a terceira da cara da distribuição de X_1. Não conseguimos calcular o limite em δ e h quando essas grandezas tendem a 0 independentemente, então vamos impor alguma relação entre eles. Defina $u := (\cos\lambda h)^{1/\delta}$. Então

$$\ln u := \frac{1}{\delta}\ln\cos(\lambda h).$$

Como estamos considerando passos h pequenos, podemos expandir o cosseno em série de Taylor em torno de h, considerando os dois primeiros termos: $\cos\lambda h \approx 1 - (\lambda^2 h^2)/2$. Agora, ainda em torno de 0, podemos expandir ln em série de Taylor considerando o primeiro termo, $\ln 1 + x \approx x$, para obter a aproximação

$$\ln u \approx \frac{1}{\delta}\ln(1 - (\lambda^2 h^2)/2) \approx -\frac{\lambda^2 h^2}{2\delta}.$$

Com isso, pela definição de u, temos

$$\mathrm{E}[\exp(i\lambda Y^{\delta,h})(t)] \approx \exp(\frac{-t\lambda^2 h^2}{2\delta}). \tag{4}$$

Escolhemos então a relação $(*)\, h^2 = \delta$ entre o tamanho do passo e tempo entre um arremesso de moeda e outro. Com isso, tomando o limite quando $\delta \to 0$ em (4), temos

$$\lim_{\delta \to 0} \mathrm{E}[\exp(i\lambda Y^{\delta,h})(t)] = \exp\left(\frac{-t\lambda^2}{2}\right). \tag{5}$$

Todo esse processo pode ser justificado de forma rigorosa, conforme (KUO, 2005, Teorema 1.2.2), que diz que há de fato um processo W_t tal que, se $(*)$ é satisfeita, então $Y_t^{\delta,h} \to W_t$ em distribuição, e a função característica de W_t é dada por (5). O caso aqui é notarmos que a função característica em (5) é justamente a função característica de uma variável aleatória com distribuição $\sim \mathcal{N}(0,t)$. Pela unicidade de funções características, somos levados a crer que o limite de passeios aleatórios que construímos nos gera um processo cuja distribuição em cada instante t é $\mathcal{N}(0,t)$. Na verdade, se formos mais a fundo, vemos que cada $Y^{\delta,h}(t) - Y^{\delta,h}(s)$ mimetiza precisamente $Y^{\delta,h}(t-s)$, propriedade básica de passeios aleatórios. Gostaríamos que a "continuização" do nosso passeio aleatório tivesse essa propriedade. Por fim, é mister notar que que cada $Y^{\delta,h}(t) - Y^{\delta,h}(s)$ é independente de $Y^{\delta,h}(u) - Y^{\delta,h}(v)$, quando $v < u < s < t$, por serem somas distintas de v.a.'s independentes. Sumarizando todas essas propriedades, se quisermos um processo que imita o passeio aleatório num ambiente contínuo, ele será algo do tipo:

Definição 0.3.1 (Movimento Browniano)**.** *Dizemos que um processo W_t em $[0,T]$ com valores em* R *é um* ***movimento Browniano unidimensional*** *iniciado em* 0 *se satisfizer:*

1. $W_0 = 0$ *qtp;*
2. $W_\omega : [0,T] \to \mathrm{R}$ *é contínuio ω-qtp.*
3. $W_t - W_s \sim \mathcal{N}(0, t-s), \qquad \forall 0 \leqslant s < t \leqslant T;$
4. *Se* $0 \leqslant t_0 < t_1 < t_2 < \ldots t_n \leqslant T$, *então os processos* $W_{t_n} - W_{t_{n-1}}, W_{t_{n-1}} - W_{t_{n-2}}, \ldots, W_{t_1} - W_{t_0}, W_{t_0}$ *são independentes.*

Um movimento Browniano em R^n *é um processo a valores em* R^n *tal que cada componente constitui um movimento Browniano unidimensional e tal que cada componente gere processos unidimensionais independentes.*

É possível mostrar a existência de movimentos Brownianos, e uma tal construção (também partindo de somas de v.a.'s i.i.d.'s), além do estudo de propriedades do movimento Browniano, são feitas de forma impecável em (STEELE, 2001, Capítulo 3) e (STEELE, 2001, Capítulos 4, 5), respectivamente. Terminamos com as seguintes observações: 1. cada caminho do movimento Browniano possui propriedades fractais. E isso é esperado pois cada pedaço de caminho, não importa quão pequeno seja, representa a oscilação do "valor total" ganho por infindáveis arremessos de moeda, a mesma propriedade que define o movimento como um todo; 2. se enxergarmos dW_t como tamanho do passo de um movimento Browniano e dt um infinitesimal de tempo onde se joga uma moeda, então a relação $(*)$ se traduz em $(dW_t)^2 = dt$. Isso nos mostra como diferenciais estocásticas se comportam de forma não usual. Estudaremos isso na seção 1.5 e as consequências disso serão diversas pelo texto; 3. a relação $(*)$ é muitas vezes chamada de **condição de difusão**.

1 Cálculo de Itô

Introduziremos, neste capítulo, a integral estocástica de Itô e desenvolveremos o cálculo relacionado a ela. Veremos como ela nasce da tentativa de se generalizar a ideia de uma transformada martingale para martingales a tempos contínuos e que, assim, ela naturalmente é uma boa representação para a riqueza de um investidor, por exemplo. Mostraremos que essa construção não pode ser feita de forma ingênua, segundo a noção já consolidada de integral de Lebesgue-Stieltjes, e como ela pode ser obtida segunda uma isometria entre espaços completos. Demonstraremos suas principais propriedades e usá-las-emos para definir as equações diferenciais estocásticas, que modelam fenômenos envolvendo aleatoriedade. Para começar o capítulo, deixaremos fixado um espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$, onde temos definida uma sequência i.i.d. de variáveis aleatórias em R, $(X_n)_{n=1}^{\infty}$, e $X_0 = 0$, tais que $\mathrm{P}(X_n = 1) = \mathrm{P}(X_n = -1) = 1/2$, $\forall n \in \mathrm{N}$. Essa sequência definirá um passeio aleatório $S_n = \sum_{i=1}^{n} X_n$ e uma filtração $\mathcal{F}_n$ em $(\Omega, \mathcal{F})$; note que a filtração comporta o conhecimento de mundo que cada jogada de moeda até n nos proporciona.

1.1 Motivação

Imagine um ativo A no mercado, cujo preço, medido nos instantes $n = 0, 1, 2, \ldots$, é dado por um passeio aleatório S_n. É como se, a cada instante n, uma moeda honesta com faces 1 e -1 fosse jogada a fim de se decidir se o preço do ativo deveria subir ou descer uma unidade. Notemos que essa modelagem não é absurda se considerarmos que, por serem infindáveis as variáveis que determinam o preço de um ativo, ao tomarmos unidades de tempo suficientemente pequenas de maneira que seu preço não varie mais que uma unidade nesses intervalos, acaba sendo razoável supor que esse preço pode subir ou descer, por quaisquer motivos que sejam, mas com igual probabilidade. Agora suponha que temos um investidor do ativo A no mercado. Esse investidor, a cada instante n, decide comprar uma determinada quantidade ξ_n desses ativos antes de sua mudança de preço. Espera então essa mudança e, assim, ganha ou perde dinheiro conforme A se valoriza ou desvaloriza. Como faz a escolha antes da mudança de preço de A, sua escolha de ξ_n é baseada apenas na informação que possui sobre o comportamento de A até o instante $n-1$. Traduzimos isso dizendo que o processo ξ_n é **previsível**, isto é,

$$\xi_n \in \mathcal{F}_{n-1}, \qquad \forall n = 1, 2, \ldots, \tag{1.1}$$

No fundo, o investidor muito se assemelha a um jogador, e sua riqueza V_n, obtida apenas através desse jogo, é

$$V_n = \xi_1(S_1 - S_0) + \xi_2(S_2 - S_1) + \ldots \xi_n(S_n - S_{n-1}). \tag{1.2}$$

Lembrando que o preço S_n do ativo A é um martingale em relação à filtração $\mathcal{F}_n$, (1.2) acima nos motiva a definir a **transformada martingale** do martingale S_n pelo processo previsível e limitado [1] ξ_n como sendo o processo $\xi \bullet S = (\xi \bullet S_n)$ dado por

$$\xi \bullet S_n = V_n = \sum_{i=1}^{n} \xi_i (S_i - S_{i-1}). \tag{1.3}$$

Note que $V_0 = 0$, e que não consideramos investimentos ξ sendo feitos no instante 0. Não podemos honestamente esperar que o investidor, organizando de maneira esperta seus investimentos (ou apostas), venha de fato a ganhar dinheiro com esse investimento. E isso porque, por mais bem planejados seus investimentos, ele não pode "olhar o futuro" de maneira a ganhar, certamente, dinheiro com seu experimento financeiro. Isso se reflete na seguinte propriedade da transformada martingale:

Proposição 1.1.1. *Seja ξ_n um processo previsível e limitado, e seja S_n um martingale. Então o processo $\xi \bullet S$ é um martingale em relação à filtração $\mathcal{F}_n$.*

Demonstração. Comece notando que se C for a constante que limita o processo ξ, então $|\xi \bullet S| \leqslant \sum_{i=1}^{n} |\xi(S_i - S_{i-1})| \leqslant C \sum_{i=1}^{n} (|S_i| + |S_{i-1}|) \leqslant 2Cn$, de forma que a transformada é integrável pra todo n. Além disso,

$$\begin{aligned} \mathrm{E}[\xi \bullet S_n | \mathcal{F}_{n-1}] &= \mathrm{E}[\sum_{i=1}^{n} \xi_i (S_i - S_{i-1}) | \mathcal{F}_{n-1}] = \sum_{i=1}^{n-1} \xi_i (S_i - S_{i-1}) \\ &+ \mathrm{E}[\xi_n (S_n - S_{n-1}) | \mathcal{F}_{n-1}] = \xi \bullet S_{n-1} + \xi_n \mathrm{E}[S_n - S_{n-1} | \mathcal{F}_{n-1}] = \xi \bullet S_{n-1}, \end{aligned}$$

onde a última igualdade vem de S_n ser martingale e a penúltima vem de ξ_n ser $\mathcal{F}_{n-1}$ mensurável. O resultado segue. □

A transformada martingale por ser re-escrita para obtermos

$$\Delta V_n = V_n - V_{n-1} = \xi_n (S_n - S_{n-1}), \tag{1.4}$$

ou seja, a variação na riqueza do investidor é dada exclusivamente pela variação no preço de A. Dizemos que uma estratégia de investimento desse tipo é *autofinanciada*. Esse tipo de estratégia será amplamente usada na seção 2.7 quando formos precificar derivativos. Vale ressaltar que a mesma introdução às transformadas martingales é dada em (CATUOGNO, 2015). Surge então a questão sobre como definir, analogamente, uma transformada martingale para processos (com especificidades que esclareceremos mais à frente) que são dados em tempo contínuo. Ou seja, intuitivamente, se o preço do ativo A variasse a todo instante de forma análoga a um passeio aleatório gerado por uma moeda

[1] Pedir que o processo ξ_n seja limitado é uma condição técnica necessária para demonstrar a propriedade de martingale da transformada martingale. Contudo, é um requerimento bem razoável em termos da modelagem, uma vez que não há "dinheiro infinito" para se investir.

jogada continuamente, portadora contudo de faces com valores proporcionalmente menores, de qual maneira poderíamos expressar a grandeza do investidor que investe também a cada instante? Lembrando que o processo de "continuização" de um passeio aleatório gera um movimento Browniano W_t (cf. 0.3), formalmente nós gostaríamos de dar sentido à expressão (advindas de (1.3) e (1.4))

$$V_t = \int_0^T dV_t = \int_0^T \xi_t dW_t. \tag{1.5}$$

A primeira tentativa de dar sentido à expressão (1.5) é numa base $\omega - \omega$. Como faríamos isso? Vamos voltar ao espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$ onde o passeio aleatório foi definido. Podemos entender Ω como o conjunto de possíveis estados, englobando cada instante n, do lance da moeda. Dessa forma, um elemento de Ω seria uma sequência de 1's e -1's, representando cada comportamento possível de um jogo de moedas em todos instantes n. Sendo assim, para cada trajetória ω do jogo de moedas, podemos calcular, explicitamente, $V_n(\omega)$, substituindo os valores $\Delta S_i(\omega)$ do passeio aleatório para aquela configuração específica de mundo, e os valores $\xi_i(\omega)$ da estratégia tomada. Pensando por esse lado, é natural tentar definir a integral em (1.5) também numa base $\omega - \omega$. Iríamos então ao espaço de probabilidade $(\Omega', \mathcal{F}', \mathrm{P}')$ onde o movimento Browniano é construído, e, para cada estado de mundo ω, tomando a trajetória de investimento $\xi_t(\omega)$, calcularíamos a integral

$$\int_0^T \xi_t(\omega) dW_t(\omega), \tag{1.6}$$

que nos forneceria a riqueza final do investidor entre os instantes 0 e T, dentro da situação ω. A questão portanto se resume a entender a integral (1.6). Sendo $\xi(\omega)$ e $W(\omega)$ funções em $[0, T]$, a integral de $\xi(\omega)$ "contra" $dW(\omega)$ deveria ser interpretada como uma integral de Lebesgue-Stieltjes. Acontece que o integrador $dW(\omega)$, para gerar um medida com sinal nos borelianos de R, precisa ter **variação finita**. De fato, temos uma medida com sinal em $[0, T]$ se, e só se, essa medida é decomponível em parte positiva e parte negativa, o que ocorre se, e só se, as funções de distribuição acumulada dessas medidas são monótonas (etc.), o que, por sua vez, ocorre se, e só se, a subtração dessas funções tem variação finita. Ou seja, reforçamos a **necessidade** de $W(\omega) : [0, T] \to \mathrm{R}$ ter variação finita a fim de se definir (1.6) como integral de Lebesgue-Stieltjes. Ver (FOLLAND, 2013, Teorema 3.27) para detalhes. Lembremos que a **variação total** de uma função $f : [a, b] \to \mathrm{R}$, denotada por $V_{a,b}f$, é dada por

$$V_{a,b}f = \sup_{\pi \in \Pi} \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|, \tag{1.7}$$

onde $\pi = \{a = t_0 < \ldots < t_n = b\}$ é uma partição de $[a, b]$, e Π é o conjunto de partições do intervalo. Temos então o seguinte resultado, retirado de (CATUOGNO, 2013, Teorema 9)

Teorema 1.1.2. *Considere o movimento Browniano W_t definido no espaço de probabilidade $(\Omega', \mathcal{F}', \mathrm{P}')$. Então, $W(\omega) : [0, T] \to \mathrm{R}$ não tem variação finita para todos $\omega \in A$, onde $\mathrm{P}(A) = 1$.*

Demonstração. Ver (CATUOGNO, 2013, Teorema 3). Ou ver (CATUOGNO, 2013, Teorema 9) para uma demonstração que não se adiante com o conceito de variação quadrática. □

E é justamente por isso que não podemos esperar definir (1.6) como uma integral de Lebesgue-Stieltjes. Precisamos então de uma nova teoria de integração com respeito a integradores advindos de processos estocásticos, em especial do movimento Browniano. É o que faremos na seção seguinte.

1.2 A Integral de Itô

Bem embora possa ser possível definir uma integral estocástica para integradores advindos de martingales contínuos quadrado integráveis quaisquer (ver (BASS, 2011, Capítulo 10)), nos ateremos apenas à integral em relação ao movimento Browniano. E isso porque, além de ser o principal exemplo de integral estocástica devido às suas aplicações, é a que possui o cálculo mais "simples" (já que provaremos que a variação quadrática do movimento Browniano é $(dW_t)^2 = dt$, como esperado), além de possuir propriedades específicas, como o teorema de representação de martingales (A.3.8), que nos permite resolver as BSDEs, introduzidas à frente. A ideia então é definir a integral de Itô de forma ingênua para uma determinada classe de processos, que chamaremos de processos simples, e estendê-la, via isometrias, para uma classe maior, dos processos adaptados. De antemão, alertamos que não podemos desejar integrar qualquer tipo de processo em relação ao integrando dW_t, pois, para isso, W_t teria que ter variação finita, o que vemos não ser verdadeiro. Ver (CATUOGNO, 2013) para detalhes. É por isso que nos concentraremos nos processos adaptados. Fixemos doravante um espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$ onde temos definido um movimento Browniano padrão, W_t, e consideremos a filtração Browniana $\mathcal{F}_t$. Denotaremos, nesta seção, por $H^2 = H^2([0, T] \times \Omega)$ o conjunto dos processos f adaptados tais que $\mathrm{E}[\int_0^T f^2(t, \omega)dt] < \infty$, e por $L^2 = L^2(\Omega)$ o conjunto das variáveis aleatórias quadrado integráveis. Seguiremos então os passos feitos em (CATUOGNO, 2013, Seção 1.1).

Definição 1.2.1. *Um processo $\phi : \Omega \times [0, T] \to \mathrm{R}$ é dito **simples** quando for do tipo*

$$\phi_t(\omega) = \sum_{i=0}^{n-1} \xi_i(\omega) 1_{(t_i, t_{i+1}]}(t),$$

onde $0 = t_0 < \ldots < t_n = T$ é uma partição do intervalo $[0, T]$ e cada ξ_i é uma variável aleatória $\mathcal{F}_{t_i}$ mensurável e limitada. Chamamos de H_0^2 o conjunto dos processos simples.

Notemos que $H_0^2 \subseteq H^2$. Para tais processos, é *natural* definirmos a **integral de Itô** como sendo a variável aleatória dada por

$$(\int_0^T \phi_t dW_t)(\omega) = \sum_{i=0}^{n-1} \phi_i(\omega)(W_{t_{i+1}}(\omega) - W_{t_i}(\omega)). \tag{1.8}$$

Também, pela limitação dos ξ_i, denotando $W_{t_{i+1}} - W_{t_i}$ por ΔW_i, temos

$$\mathrm{E}[(\int_0^T \xi_t dW_t)^2] = \mathrm{E}[(\sum_{i=0}^{n-1} \xi_i(W_{t_{i+1}} - W_{t_i}))^2] = \mathrm{E}(\sum_{i,j=0}^{n-1} \xi_i\xi_j \Delta W_i \Delta W_j) \leqslant CT, \tag{1.9}$$

onde usamos a independência de incrementos do movimento para calcular $\mathrm{E}[\Delta W_i \Delta W_j] = \delta_{i,j}(t_{i+1} - t_i))$. Assim, para cada processo simples $\phi \in H_0^2$, a integral $\int_0^T \phi dW_t$ é uma variável aleatória em L^2, sendo ela então um mapa entre os espaços normados H_0^2 e L^2. Da definição, é fácil ver a primeira propriedade da integral de Itô.

Proposição 1.2.2. *H_0^2 é subespaço vetorial de H_2 e a integral é linear como mapa de H_0^2 a L^2.*

Além disso, este mapa é uma isometria, como visto no próximo teorema, devido a Itô.

Teorema 1.2.3 (Isometria de Itô.)**.** *Para todo processo simples $\phi \in H_0^2$, temos*

$$\mathrm{E}[\int_0^T \phi^2 dt] = \mathrm{E}[(\int_0^T \phi dW_t)^2]. \tag{1.10}$$

Demonstração. Tome ϕ como na definição 1.2.1.Comecemos calculando a integral da direita. Para isso, aproveitamos (1.9):

$$\mathrm{E}[(\int_0^T \phi dW_t)^2] = \sum_{i,j}^{n} \mathrm{E}[\xi_i\xi_j \Delta W_i \Delta W_j] = \sum_{i=1}^{n} \mathrm{E}[\xi_i^2](t_{i+1} - t_i) + \sum_{i \neq j} \mathrm{E}[\xi_i\xi_j \Delta W_i \Delta W_j]. \tag{1.11}$$

Agora, note que se $i < j$, por exemplo,

$$\mathrm{E}[\xi_i\xi_j \Delta W_i \Delta W_j] = \mathrm{E}[\mathrm{E}[\xi_i\xi_j \Delta W_i \Delta W_j \mid \mathcal{F}_j] =$$
$$\mathrm{E}[\xi_i\xi_j \Delta W_i \mathrm{E}[\Delta W_j \mid \mathcal{F}_j]] = \mathrm{E}[\xi_i\xi_j \Delta W_i \mathrm{E}[\Delta W_j]] = 0,$$

onde usamos a independência de incrementos do movimento Browniano. Substituindo em (1.11),

$$\mathrm{E}[(\int_0^T \phi dW_t)^2] = \sum_{i=1}^{n} \mathrm{E}[\xi_i^2](t_{i+1} - t_i) = \mathrm{E}[\sum_{i=1}^{n} \xi_i^2 (t_{i+1} - t_i)] = \mathrm{E}[\int_0^T \phi^2 dt],$$

onde usamos que $\phi^2(\omega, t) = \sum_{i=0}^{n-1} \xi_i^2 1_{(t_i, t_{i+i}]}$ e integramos para cada ω normalmente segundo Lebesgue. □

Com isso, somos levados a estender a noção de integral de Itô para o fecho de H_0^2 em H^2. Antes disso, notamos que podemos entender a integral também como um processo, se definirmos de forma óbvia, para $\phi \in H_0^2$,

$$\int_0^t \phi dW_t := \int_0^T 1_{[0,t]} \phi dW_t. \tag{1.12}$$

O processo assim definido satisfaz algumas propriedades:

Proposição 1.2.4. *Seja $\phi \in H_2^0$. Então:*

1. *A integral de Itô $\{\int_0^t \phi_s dW_s\}_{t=0}^T$ é contínua.*
2. *A integral de Itô $\{\int_0^t \phi_s dW_s\}_{t=0}^T$ é um martingale quadrado integrável.*

Demonstração. Note que a continuidade segue claramente da continuidade do movimento Browniano. A outra parte é uma conta simples encontrada em (CATUOGNO, 2013, Teorema 5). □

Agora, vemos até onde podemos estender a integral.

Proposição 1.2.5. *O espaço H_0^2 é denso em H^2.*

Demonstração. A demonstração é simples e consiste apenas em notar que a sequência de processos simples

$$\phi_n(\omega, t) := \sum_{i=1}^{2^n - 1} \Big[\frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_t} \phi(\omega, u) du \Big] 1_{(t_i, t_{i+1}]},$$

onde $t_i = iT/2^n$, $i = 0, 1, \ldots, 2^n$, aproxima ϕ em H^2. Os detalhes estão bem apresentados em (STEELE, 2001, Teorema 6.5). □

Seja $f \in H^2$ e tome uma sequência ϕ^n de processo em H_0^2 que aproxima f na norma $|| \cdot ||_{H^2}$. Note que ϕ^n é uma sequência de Cauchy em H^2. Dessa forma, as variáveis aleatória $I_n(f) = \int_0^T \phi_t^n dW_t$ formam também uma sequência de Cauchy em L^2, já que

$$||I_m(f) - I_n(f)||_{L^2} = || \int_0^T \phi_t^m dW_t - \int_0^T \phi_t^n dW_t||_{L_2} =$$
$$|| \int_0^T (\phi_t^m - \phi_t^n) dW_T||_{L_2} = ||\phi_m - \phi_n||_{H^2},$$

pela linearidade da integral e pela isometria de Itô. Com isso, sendo L^2 completo, existe uma única variável aleatória em L^2 $I(f, \phi^n)$ tal que $I_m(f) \to I(f, \phi^n)$. O próximo lema mostra que a variável assim obtida não depende da escolha da sequencia ϕ_n em H_0^2.

Lema 1.2.6. *Se (ϕ^n) e (ψ^n) são sequências em H_2^0 qu tendem a f em na norma $||\cdot||_{H^2}$, então $I(f, \phi^n) = I(f, \psi^n) := I(f)$*

Demonstração. Isso na verdade é só um caso particular do resultado mais geral de espaços métricos: se M e N são espaços métricos com N completo, $f : M \to N$ uma isometria entre eles, e (x_n) e (y_n) duas sequências em M que tendem a a, então $f(x_n)$ e $f(y_n)$ são duas sequências em N que tendem ambas a $f(a)$. □

Definimos, dessa forma, para qualquer processo $f \in H$, sua integral de Itô por $I(f)$. De maneira análoga, podemos definir o **processo** $(I_t(f))$ como sendo $I(f 1_{[0,t]})$. A questão é que as integrais de Itô são dadas em L^2 de tal maneira que elas podem ser determinadas ambiguamente em conjuntos de medida nula. Isso significa que, para cada $t \in [0, T]$, a integral estocástica $(I_t(f))$, que é uma variável aleatória dada por um limite de sequência em L^2, pode não ser bem definida num conjunto de medida nula, N_t. Ao considerarmos todas as integrais $I_t(f)$, com t percorrendo T, pode ser que $\mathrm{P}(\bigcup_{t \in [0,T]} N_t) \neq 0$, e assim não teríamos o processo $(I_t(f))$ bem definido em um conjunto considerável. Há, contudo, um teorema que resolve esse problema:

Teorema 1.2.7. *Para todo processo $f \in H^2$, há um processo contínuo X_t tal que $\mathrm{P}(X_t = I_t(f)) = 1$, para todo $t \in [0, T]$.*

Demonstração. Ver (STEELE, 2001, Teorema 6.2). □

Consideramos então, sempre, essa versão contínua como sendo a integral de Itô do processo em questão, já que continuidade é uma propriedade padrão que desejamos em integrais. E ela satisfaz a seguinte propriedade, analogamente às transformadas martingales:

Proposição 1.2.8. *A integral $I_t(f)$ de um processo $f \in H^2$ é um martingale quadrado integrável.*

Demonstração. Não deve ser ter dúvidas sobre a quadrado integrabilidade do processo, pela fórmula de Itô. Para o resto, ver novamente (STEELE, 2001, Teorema 6.2). □

Notemos que a isometria de Itô vale também para qualquer processo adaptado, ou seja,

$$||f||_{H^2} = ||I(f)||_{L^2} \; \forall f \in H^2, \tag{1.13}$$

já que

$$||I(f)||_{L^2} = \lim_n ||I(\phi^n)||_{L^2} = \lim_n ||\phi||_{H^2} = ||f||_{H^2},$$

pela isometria de Itô para processos simples.

Antes de encerrar a seção, observamos que enquanto a transformada martingale pode ser compreendida numa base $\omega-\omega$ para além do costumeiro estudo de sua distribuição de probabilidades, o mesmo não vale para a integral de Itô, a princípio. Na verdade, ela é definida numa base $\omega - \omega$ para processos simples, porém sua extensão se dá via isometrias. Nos perguntamos se há alguma maneira de compreender melhor os caminhos para um processo definido via integral de Itô. Esse é o conteúdo do próximo teorema.

Teorema 1.2.9. *Seja $f \in H^2$ e seja τ um $\mathcal{F}_t$ - tempo de parada tal que $f(s,\omega) = 0$ para quase todos ω em $\{\omega | s \leqslant \tau(\omega)\}$. Então*

$$\int_0^t f(\omega, s)dW_s = 0, \tag{1.14}$$

para quase todos $\omega \in \{\omega | t \leqslant \tau(\omega)\}$.

Demonstração. Ver (STEELE, 2001, Teoremas 6.3 e 6.4). □

1.3 Uma Extensão para a Integral de Itô

Nesta seção, seguiremos lado a lado a construção da integral de Itô para processos localmente H^2 feita em (STEELE, 2001, capítulo 7). Lembramos que um processo X em $[0, T]$ é dito satisfazer a propriedade P localmente se existir uma sequência não decrescente de tempos de parada τ_n em $[0, T]$ tal que

1. $\mathrm{P}(\bigcup_{n=1}^{\infty} \{\tau_n = T\}) = 1$;
2. $X_t^{\tau_n} := X_{t \wedge \tau_n}$ satisfaz P.

Dizemos então que um processo f é **localmente** H^2 se existe uma sequência não decrescente de tempos de parada τ_n satisfazendo (1) acima e tal que

$$f_n(\omega, t) := f(\omega, t) 1_{\{t \leqslant \tau_n(\omega)\}} \in H^2, \qquad \forall n. \tag{1.15}$$

A questão então é porque tentar estender a integral de Itô a esse tipo de processo. Definimos o conjunto H^2_{loc} como sendo o conjunto dos processos f_t tais que

$$\mathrm{P}(\omega | \{\int_0^T f^2(\omega, t)dt < \infty\}) = 1, \tag{1.16}$$

e temos o seguinte resultado:

Proposição 1.3.1. *Um processo f é localmente H^2 se, e só se, $f \in H^2_{loc}$*

Demonstração. Suponha que f seja localmente H^2, e tome uma sequência de tempos de parada como acima. Então, existe $\Omega_0 \subseteq \Omega$ tal que $\mathrm{P}(\Omega_0) = 1$ e, para todo $\omega \in \Omega_0$, existe n tal que $\tau_n(\omega) = T$. Ainda, por hipótese, cada $f_n \in H^2$, ou seja, $\mathrm{E}[\int_0^T f_n^2 dt] < \infty$, de forma que existe uma sequência de subconjuntos mensuráveis Ω_i, $i = 1, 2, 3, \ldots$ tal que $\mathrm{P}(\Omega_i) = 1$ para todo i e $\int_0^T f_n^2(\omega)dt < \infty$ para todo $\omega \in \Omega_n$. Considere agora o conjunto mensurável $\Omega' := \bigcap_{i=0}^{\infty} \Omega_i$. Note que $\mathrm{P}(\Omega') = 1$ e, se $\omega \in \Omega'$, então $\tau_n(\omega) = T$ para algum n, de maneira que $f_n(\omega, t) = f(\omega, t)$ para todo t e, portanto, $\int_0^T f^2(\omega, t)dt = \int_0^T f_n^2(\omega, t)dt < \infty$, já que $\omega \in \Omega_n$. Assim, $\mathrm{P}(\omega|\{\int_0^T f^2(\omega, t)dt < \infty\}) \geqslant \mathrm{P}(\Omega') = 1$ e $f \in H^2_{loc}$.

Agora, seja $f \in H^2_{loc}$. Considere a sequência de tempos de parada:

$$\tau_n(\omega) := \inf\{s : \int_0^s f(\omega, t)dt \geqslant n \,\text{ou}\, s \geqslant T\}. \tag{1.17}$$

Note que é uma sequência não decrescente e que existe Ω_0 tal que $\mathrm{P}(\Omega_0) = 1$ e, se $\omega \in \Omega_0$, então $\int_0^T f^2(\omega, t)dt < \infty$, ou seja, existe n tal que $\int_0^T f^2(\omega, t)dt < n$. Dessa forma, $\tau_n(\omega) = T$. Portanto, $\mathrm{P}(\bigcup_{n=1}^{\infty}\{\tau_n = T\}) = 1$. Ainda, pela definição dos τ_n, temos $\int_0^T f(\omega, t)1_{\{t \leqslant \tau_n(\omega)\}}dt \leqslant n \Rightarrow f_n \in H^2$. Assim, mostramos que f é localmente H^2, e o resultado segue. □

Temos então uma equivalência entre processos localmente quadrado integráveis e processos em H^2_{loc}. Este último conjunto é de interesse pois é claramente mais abrangente que H^2, a ponto de conter todos os processos do tipo $f(W_t)$, com f uma função contínua. De fato, neste caso, para quase todos ω, a função $f(W_t)^2$ é contínua e, portanto, limitada, o que faz que $f(W_t)$ verifique facilmente (1.16). E é extremamente desejável, numa teoria de integração em relação ao integrador dW_t, conseguirmos integrar processos do tipo $f(W_t)$, com f contínuo. Para $f = \mathrm{Id}$ isso já ocorre, uma vez que $W_t \in H^2$ (e $\int_0^T W_t dW_t = W_t^2/2 - t/2$, como veremos, o que nos confirma que o segundo processo é um martingale). Claro que surge o questionamento sobre não ser óbvio o pertencimento de qualquer processo do tipo $f(W_t)$, f contínuo, a H^2. A resposta é negativa, pelo exemplo abaixo.

Exemplo 1.3.2. *O processo* $\exp(W_t^4)$ *não pertence a* H^2*. De fato,*

$$\mathrm{E}[\int_0^T (\exp(W_t^4))^2 dt] = \int_0^T \mathrm{E}[\exp(2W_t^4)]dt. \tag{1.18}$$

Lembrando que $W_t \sim \mathcal{N}(0,t)$*, para* $t > 0$*, temos (para* $t > 0$ *fixado):*

$$\mathrm{E}[\exp(2W_t^4)] = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{2x^4} e^{-\frac{x^2}{2t}} dx = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{(\sqrt{2}x^2 - \frac{1}{4\sqrt{2}t})^2} e^{\frac{1}{32t^2}} dx,$$

e essa ultima integral não converge, já que o integrando $\to \infty$ *quando* $x \to \infty$*. Dessa forma, a norma* L^2 (1.18) *do processo em questão não é finita.*

Com isso, nasce a necessidade de uma extensão da noção de integração para abranger estes processos, e a construiremos a seguir. Comecemos notando que a sequência de tempos de parada em (1.17) é um exemplo de sequência que **localiza** processos $f \in H^2_{loc}$ para processes $f_n \in H^2$. Esses últimos possuem integral de Itô bem definida, com versão contínua $I_t(f_n)$. A ideia então é mostrar que essa sequência de integrais de Itô converge em um determinado sentido, e definir esse limite como a integral de f.

Proposição 1.3.3. *Para qualquer* $f \in H^2_{loc}$*, dada uma sequência de tempos de parada* τ_n *que localiza* f *em* H^2*, se* $X_{t,n}$ *é o processo contínuo que representa integral de Itô de* f_n*, então, para qualquer* $t \in [0,T]$ *e* $n \geqslant m$

$$X_{t,n} = X_{t,m} \text{ para quase todos } \omega \in \{\tau_m(\omega) \geqslant t\}$$

Demonstração. Note que as funções $f_n(\omega,t) = f(\omega,t)1_{\{t\leqslant\tau_n(\omega)\}}$ e $f_m(\omega,t) = f(\omega,t)1_{\{t\leqslant\tau_m(\omega)\}}$ são iguais no conjunto $\{\omega|\tau_m(\omega) \geqslant t\}$, já que $\tau_m \leqslant \tau_n$. O resultado segue como corolário direto do teorema 1.2.9. □

Essa proposição nos diz que, fixado um tempo t e um índice m, se tomarmos apenas os ω que são parados por τ_m após o instante t, a sequência $X_{t,m+p}(\omega)$, $p = 0,1,2,\dots$ se torna constante $:= c(t,\omega)$. Isso nos induz a definirmos a integral de Itô de f no instante t, para ω's nesse conjunto, por $c(t,\omega)$. Como $\tau_n(\omega) \nearrow T$ qtp, eventualmente conseguimos definir um valor para a integral de Itô para todos os pares (ω,t), como desejamos, dessa forma intuitiva. Contudo, não nos é garantido a continuidade de um processo integral definido dessa forma. E é com isso que lidaremos no próximo resultado.

Teorema 1.3.4. *Tomando* $X_{t,n}$ *como na proposição 1.3.3, temos que* ***existe*** *um processo contínuo* X_t *tal que*

$$\mathrm{P}(X_t = \lim_{n\to\infty} X_{t,n}) = 1 \qquad \forall t \in [0,T].$$

Demonstração. Defina a variável aleatória

$$N(\omega) := \min\{n : \tau_n(\omega) = T\}.$$

Fazendo a nomeação $\Omega_0 := \{\omega|N(\omega) < \infty\}$, pela definição da sequência localizadora, temos $\mathrm{P}(\Omega_0) = 1$. Assim, para cada $\omega \in \Omega_0$, definimos o processo

$$X(t,\omega) = X_{t,N(\omega)}(\omega).$$

Pela continuidade de cada elemento da sequência $X_{t,n}$, é claro que X_t é contínuo também. Ainda, temos que $\Omega_0 = \bigsqcup_{i=1}^{\infty} \Omega_i$, onde $\Omega_i = \{\omega | N(\omega) = i\}$. Defina $A = A(t) = \{\omega | \lim_{n\to\infty} X_{t,n}(\omega) = X_t(\omega) = X_{t,N(\omega)}(\omega)\}$. Mostraremos que $\mathrm{P}(A) = 1$ para todo $t \in [0, T]$.

$$\mathrm{P}(A) = \mathrm{P}(\bigsqcup_{i=1}^{\infty}(A \cap \Omega_i)) = \sum_{i=1}^{\infty} \mathrm{P}(A \cap \Omega_i) = \sum_{i=1}^{\infty} \mathrm{P}(A|\Omega_i)\mathrm{P}(\Omega_i). \tag{1.19}$$

Considerando Ω_i com a probabilidade induzida, vemos que $X_{t,i+p} = X_{t,i+p+1}$ com probabilidade 1, para todo $p \in \mathbb{N}$, pela proposição anterior. Logo, também com probabilidade 1, $X_{t,i+p}(\omega)$ é constante e igual a $X_{t,i}(\omega) = X_{t,N(\omega)}(\omega)$, em Ω_i. Assim, $\mathrm{P}(A|\Omega_i) = 1$ para todo i. Logo, (1.19) fica

$$\mathrm{P}(A) = \sum_{i=1}^{\infty} \mathrm{P}(\Omega_i) = \mathrm{P}(\Omega_0) = 1,$$

e o resultado segue. □

Definimos, assim, para um processo f localmente H^2, a sua integral de Itô como o processo X_t construído pelo teorema 1.3.4. Surgem então duas perguntas naturais: o quanto essa construção de integral de Itô para $f \in H^2_{loc}$ depende da sequência localizadora; e se essa noção de integral coincide com a noção anterior, para processos em H^2, o que se espera ocorrer, já que queremos uma **extensão**. Responderemos com o próximo teorema e um corolário óbvio.

Teorema 1.3.5. *Sejam τ_n e τ'_n duas sequências que localizam o processo $f \in H^2_{loc}$. Sejam $X_{t,n}$ e $X'_{t,n}$ as integrais de Itô de cada f_n e f'_n, definidas de forma óbvia. Então,*

$$\mathrm{P}(\lim_{n\to\infty} X_{t,n} = \lim_{n\to\infty} X'_{t,n}) = 1.$$

Demonstração. Ver (STEELE, 2001, Proposição 7.4). □

Corolário 1.3.6. *Seja $f \in H^2$. Então, a integral de Itô de f definida de forma usual, $I(f)_t$, coincide com a integral de Itô X_t definida por meio do limite das integrais $X_{t,n}$ de f localizada.*

Demonstração. Considere a sequência $\tau_n = T$ e note que ela claramente localiza f. Ainda, $f_n = f$ para todo n, de forma que $X_{t,n} = I(f_n)_t = I(f)_t$. Como a integral X_t não depende da localização escolhida, o resultado segue. □

De agora em diante, denotaremos de $I_t(f)$ a integral de Itô de um processo $f \in H^2_{loc}$ qualquer, como definida acima. Antes de encerrar a seção, vale notar que para esses processos não é necessariamente verdade que $I_t(f)$ seja um martingale. Contudo, é verdade que $I_t(f)$ é um martingale local. Basta tomar uma sequência localizadora τ_n qualquer para f e notar que $I_{t\wedge\tau_n}(f) = I_t(f_n)$, com $f_n \in H^2$, e essa última integral é um martingale, como queríamos.

1.4 A Fórmula de Itô

De posse de uma nova teoria de integração, gostaríamos de entender o cálculo associado. Vale notar que trabalharemos com processos contínuos. Para o cálculo usual de funções $f : \mathrm{R} \to \mathrm{R}$, isso se resume a entender as duas componentes do Teorema Fundamental do Cálculo e usá-las para computar derivadas e integrais:

Parte 1. Se $f : [a, b] \to \mathrm{R}$ é integrável, então $F : [a, b] \to \mathrm{R}$, $x \mapsto \int_a^x f(t)dt$ é diferenciável e $F'(x) = f(x)$.

Parte 2. Se $F : [a, b] \to \mathrm{R}$ é continuamente diferenciável, então

$$\int_a^b F'(x)dx = F(b) - F(a). \tag{1.20}$$

Notemos que não podemos esperar encontrar uma transcrição ingênua da parte 1 para o cálculo estocástico, uma vez que mesmo a mais simples das integrais, $\int_0^t 1dW_t = W_t$, não é derivável (já que o movimento Browniano possui variação ilimitada qtp). Tentemos então dar significado estocástico à parte 2. Faremos isso, na verdade, com uma modificação de (1.20). Sejam f e g funções de R em R diferenciáveis. Então, a regra da cadeia nos diz que

$$\frac{d}{dx}f(g(x)) = f'(g(x))g'(x) \Rightarrow f(g(x)) - f(g(0)) = \int_0^x f'(g(x))g'(x)dx, \tag{1.21}$$

onde usamos (1.20) na implicação. Quando g é de variação limitada, por exemplo quando g' é limitada, (1.21) acima pode ser escrita como

$$f(g(x)) - f(g(0)) = \int_0^x f'(g(x))dg(x), \tag{1.22}$$

onde consideramos a medida de Lebesgue-Stieltjes associada a g. A fórmula acima, muitas vezes abreviada por

$$d(f \circ g) = f' \circ g dg, \tag{1.23}$$

é a que procuraremos adaptar ao caso de integrais estocásticas, e veremos que pode ser utilizada para encontrar concretamente algumas integrais. Comecemos com uma definição.

Definição 1.4.1. *Dizemos que um processo $X : \Omega \times [0, T] \to \mathrm{R}$ é um processo de Itô (simples e unidimensional) quando X tiver a representação integral*

$$X(t, \omega) = x + \int_0^t a(s, \omega)ds + \int_0^t b(s, \omega)dW_s,$$

onde a e b são processos progressivamente mensuráveis tais que

$$\mathrm{P}(\int_0^T |a(s, \omega)|ds < \infty) = 1, \tag{1.24}$$

e

$$\mathrm{P}(\int_0^T |b(\omega, s)|^2 ds < \infty) = 1. \tag{1.25}$$

Neste caso, escrevemos também

$$dX_t = a_t dt + b_t dW_t$$

A ideia então é, dado um processo e Itô X_t e uma função suficientemente suave f, escrever $f(X_t)$ como um processo de Itô, similarmente ao feito em (1.22). Para tanto, precisamos de um resultado padrão em análise estocástica, conhecido como decomposição de Doob-Meyer, e cuja demonstração pode ser encontrada, para um resultado um pouco mais geral, em (BASS, 2011, Teorema 9.12). Notemos que estamos considerando processos em $[0, T]$ definidos num espaço de probabilidade Ω munido de uma filtração Browniana (não que ser Browniana venha ao caso para a seguinte decomposição).

Teorema 1.4.2 (Doob-Meyer decomposition)**.** *Seja Z_t um submartingale contínuo. Então, existe um processo adaptado crescente e contínuo qtp, A_t, iniciado em 0, e um martingale contínuo M_t, tais que*

$$Z_t = M_t + A_t.$$

Mais ainda, se A' e M' são outros tais processos, então $M_t = M'_t$ e $A_t = A'_t$ para todo t qtp.

Notemos que, na referência citada, o teorema é demonstrado para supermartingales contínuos de classe D definidos em $[0, \infty)$. O fato de tratarmos o caso de submartingales não é problemático, pois, para traduzir para nosso teorema, basta tomarmos o processo $-Z_t$. Ainda, não pedimos a hipótese de o processo ser de classe D por dois motivos: 1. Ela é de fato desnecessária, como podemos ver em (LOWTHER, 2011b); 2. A decomposição de Doob-Meyer é utilizada neste trabalho apenas nos casos em que Z_t é um submartingale não negativo definido em $[0, T]$, na construção de variações quadráticas. Acontece que nestes casos a condição de Z_t se de classe D é trivialmente satisfeita. Esclareceremos essa questão a seguir por completude. Eis a definição da referência sobre processos de classe D:

Definição 1.4.3. *Um processo Z (em $[0, \infty)$) é de classe D se o conjunto*

$$\{Z_\tau | \tau \text{ é um tempo de parada finito}\}$$

é uma família de variáveis aleatórias uniformemente integrável.

Recordemos que a família $\{Z_i\}_{i \in I}$ é dita uniformemente integrável quando

$$\lim_{x\to\infty} \sup_{i\in I} \mathrm{E}[|X_i|; \{|X_i| \geqslant x\}] \xrightarrow{x\to\infty} 0.$$

Temos ainda o lema:

Lema 1.4.4. *Seja a família de v.a.'s dadas por esperanças condicionais* $\{\mathrm{E}[Z|\mathcal{G}]\}_{\mathcal{G}\in G}$*, onde cada* $\mathcal{G}$ *é uma sub-*σ*-álgebra de* $\mathcal{F}$*. Então essa família é uniformemente integrável.*

Demonstração.

$$\begin{aligned} &\mathrm{E}[|\mathrm{E}[Z \mid \mathcal{G}]|;\{|\mathrm{E}[Z \mid \mathcal{G}]| \ \geqslant x\}] \leqslant \mathrm{E}[\mathrm{E}[|Z| \mid \mathcal{G}]|;\{|\mathrm{E}[Z \mid \mathcal{G}]| \geqslant x\}] \\ &= \mathrm{E}[|Z|;\{|\mathrm{E}[Z \mid \mathcal{G}]| \geqslant x\}]. \end{aligned} \tag{1.26}$$

onde a desigualdade segue de Jensen. O conjunto $\{|\mathrm{E}[Z \mid \mathcal{G}]| \geqslant x\}$ tem sua probabilidade majorada por Markov:

$$\mathrm{P}(\{|\mathrm{E}[Z \mid \mathcal{G}]| \geqslant x\}) \leqslant \frac{\mathrm{E}[\mathrm{E}[|Z| \mid \mathcal{G}]]}{x} = \frac{\mathrm{E}[|Z|]}{x}. \tag{1.27}$$

Agora, como Z é uniformemente integrável por estar em L^1, fazendo $x \to \infty$ em (1.27) nós fazemos com que a primeira esperança em (1.26) $\to 0$. O resultado segue. □

Voltando ao nosso caso, como Z_t é um submartingale, pelo teorema de parada de Doob A.3.2, temos que $Z_\tau \leqslant \mathrm{E}[Z_T \mid \mathcal{F}_\tau]$, para qualquer tempo de parada. Como estamos supondo Z não negativo, isso se traduz para $|Z_\tau| \leqslant |\mathrm{E}[Z_T \mid \mathcal{F}_\tau]| \leqslant |\mathrm{E}[|Z_T| \mid \mathcal{F}_\tau]|$. Acontece a família de v.a.'s $\{\mathrm{E}[|Z_T| \mid \mathcal{F}_\tau]\}$ é uniformemente integrável pelo lema 1.4.4. Também, quando $\{X_i\}_{i\in I}$ é uniformemente integrável e $\{Z_i\}_{i\in I}$ é tal que $|Z_i| \leqslant |X_i|$, então $|Z_i|$ também é uniformemente integrável. Com efeito,

$$\mathrm{E}[|Z_i|;\{|Z_i| \geqslant x\}] \leqslant \mathrm{E}[|X_i|;\{|Z_i| \geqslant x\}] \leqslant \mathrm{E}[|X_i|;\{|X_i| \geqslant x\}],$$

donde a afirmação segue claramente, tomando supremos em $i \in I$ e depois limite em $x \to \infty$. Assim, mostramos que a condição ser de classe D é trivialmente satisfeita para nós, e podemos seguir para a decomposição de Doob-Meyer, de onde surgem duas definições:

1. Um **semimartingale** é um processo do tipo

$$X_t = M_t + A_t,$$

onde M_t é um martingale local e e A_t é um processo de variação finita. Isso generaliza a noção de submartingales. Notemos aqui que: 1. processos de Itô são tipos especiais de semimartingales, onde a integral usual é o processo de variação limitada e a integral de Itô o martingale local; 2. a decomposição de um semimartingale é única, uma vez que martingales que possuem variação finita são nulos (ver (BASS, 2011, teorema 9.7)).

2. Pela desigualdade de Jensen, se M_t é um martingale local contínuo e quadrado integrável, então M_t^2 é um submartingale local, de forma que existe A_t como em 1.4.2, isto é, contínuo, adaptado, crescente e iniciado em 0, tal que

$$N_t := M_t^2 - A_t$$

é um martingale local. Chamamos $\langle M\rangle_t := A_t$ de **variação quadrática** de M_t. Note que, sendo a variação quadrática um processo crescente e contínuo, ela gera uma medida de Lebesgue-Stieltjes em $[0,T]$. Dados dois martingales locais, contínuos e quadrado integráveis M_t e N_t, definimos também sua **covariação** ou **variação quadrática cruzada (conjunta)** por

$$\langle M,N\rangle_t := \frac{1}{2}[\langle M+N\rangle_t - \langle M\rangle_t - \langle N\rangle_t]. \tag{1.28}$$

Chamamos a também identidade definidora (1.28) acima de **fórmula de polarização**. A variação quadrática cruzada tem a seguinte definição equivalente:

Definição 1.4.5. *Se M_t e N_t são martingales locais, contínuos e quadrado integráveis, $\langle M,N\rangle_t$ é o único processo contínuo de variação limitada tal que $M_tN_t - \langle M,N\rangle_t$ é um martingale local.*

A equivalência entre essas definições é estudada em (CATUOGNO, 2013, seção 1.6). Ainda, temos a seguinte generalização: se X_t e Y_t forem semimartingales (como processos de Itô por exemplo), definimos

$$\begin{aligned}\langle X\rangle_t &:= \langle M\rangle_t;\\ \langle X,Y\rangle_t &:= \langle M,N\rangle_t,\end{aligned} \tag{1.29}$$

onde temos as decomposições $X_t = M_t + A_t$, $Y_t = N_t + B_t$, com M e N martingales locais e A_t e B_t processos de variação limitada.

Perceba que fizemos o tratamento da variação quadrática usando sempre processos que satisfazem alguma propriedade local; em contrapartida, 1.4.2 versa sobre uma decomposição para *submartigales*, sem usar o termo local. Acontece que esse teorema pode ser facilmente estendido, por meio de localização, para abranger a hipótese de Z_t ser *submartingale local, localmente de classe D*, fornecendo agora um *martingale local* na decomposição. Podemos ver detalhes em (LOWTHER, 2011b). Estamos agora em condições de entender o teorema fundamental do cálculo de Itô, também chamado de fórmula de Itô:

Teorema 1.4.6 (Fórmula de Itô)**.** *Seja X_t um processo de Itô, e seja $f \in C^2$. Então,*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s. \tag{1.30}$$

(1.30) *também é escrita como*

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)d\langle X\rangle_t. \tag{1.31}$$

Demonstração. Ver (BASS, 2011, Teorema 11.1) □

Note a semelhança entre (1.23) e (1.31), com o acréscimo do termo $\frac{1}{2}f''(X_t)d\langle X\rangle_t$ nessa última equação. Ele surge justamente do fato de a "diferencial $dW_t = \sqrt{dt}$" não se comportar como uma diferencial usual.

Analogamente aos processos de Itô unidimensionais, temos os processos de Itô mais gerais, em relação a vários movimentos Browniano independentes. São então processos do tipo (ainda unidimensionais) $X : \Omega \times [0,T] \to \mathrm{R}$,

$$dX_t = a(t)dt + b^1(t)dW_t^1 + \ldots + b^n(t)dW_t^n,$$

onde $a, b^1 \ldots b^n$ satisfazem, respectivamente, as condições (1.24) e (1.25). Temos também uma versão multidimensional, mais geral, cuja demonstração é feita de forma similar à versão unidimensional.

Teorema 1.4.7 (Fórmula de Itô Multidimensional)**.** *Sejam $X_t^1, \ldots, X_t^d$ processos de Itô, e seja $f \in C^2(\mathrm{R}^d)$. Então, denotando $X_t = (X_t^1, \ldots, X_t^d)$, temos*

$$f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \partial_i f(X_s)dX_s^i + \frac{1}{2}\sum_{i,j=1}^{d} \int_0^t \partial_{i,j} f(X_s)d\langle X^i, X^j\rangle_s, \forall t \in [0,T], qtp. \quad (1.32)$$

Note que podemos escrever, de forma simplificada, (1.32) *como*

$$df(X_t) = \sum_{i=1}^{d} \partial_i f(X_t)dX_t^i + \frac{1}{2}\sum_{i,j=1}^{d} \partial_{i,j} f(X_t)d\langle X^i, X^j\rangle_t, \forall t \in [0,T], qtp. \quad (1.33)$$

Ressaltemos que essas duas versões da fórmula de Itô são introduzidas de forma muito similar e mais completa em (BASS, 2011, Capítulo 11).

1.5 Variação Quadrática

Notemos que ao aplicarmos a fórmula de Itô para um processo do tipo $f(X_t)$, X_t um processo de Itô d-dimensional, obtemos integrais em relação a dois tipos de integradores: dX_t^i e $d\langle X^i, X^j\rangle_t$. Por definição, dX_t^i é do tipo $a^i dt + b^{i,1}dW_t^1 + \ldots + b^{i,n}dW_t^n$, e calculamos uma integral do tipo $\int_0^t g(s)dX_s^i$ da forma óbvia (que em verdade vem da definição de integral em relação a um semimartingale)

$$\int_0^t g(s)a_s^i ds + \int_0^t g(s)b_s^{i,1}dW_s^1 + \ldots + \int_0^t g(s)b_s^{i,n}dW_s^n.$$

A questão então é encontrar os processos de variação limitada qtp $\langle X_t^i, X_t^j\rangle$. Comecemos com dois teoremas e seus corolários que nos ensinam a calcular explicitamente variações quadráticas.

Teorema 1.5.1. *Seja M um martingale local, contínuo e quadrado integrável em $[0, T]$. Então*

$$\langle M \rangle_t = \operatorname{p}\lim_{n \to \infty} \sum_{i=1}^{n} (M_{t_i} - M_{t_{i-1}})^2, \tag{1.34}$$

onde o limite em probabilidade acima se dá quando percorremos uma sequência qualquer de partições π_n de $[0, t]$ cujo tamanho $|\pi_n| \to 0$.

Demonstração. Primeiramente, precisamos mostrar que o limite em probabilidade de (1.34) existe e não depende da partição escolhida. Esse é o conteúdo de (CATUOGNO, 2013, lema 21). Que o processo definido pelo limite em probabilidade é crescente é bem claro. Por fim, que $M^2 - L$ onde L é o limite em (1.34) é uma martingale local é o conteúdo de (CATUOGNO, 2013, lema 19). Portanto, pela unicidade da decomposição de Doob-Meyer, obtemos que esse limite é de fato a variação quadrática de M. □

Corolário 1.5.2. *Seja X um semimartingale contínuo. Então,*

$$\langle X \rangle_t = \operatorname{p}\lim_{n \to \infty} \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2. \tag{1.35}$$

Demonstração. Seja $X = M + A$ a decomposição do semimartingale em martingale local e processo de variação limitada, contínuos. Temos entao que, para cada partição π,

$$\begin{aligned} \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2 &= \sum_{i=1}^{n} (M_{t_i} - M_{t_{i-1}})^2 + \\ \sum_{i=1}^{n} (A_{t_i} - A_{t_{i-1}})^2 &+ \sum_{i=1}^{n} (M_{t_i} - M_{t_{i-1}})(A_{t_i} - A_{t_{i-1}}). \end{aligned} \tag{1.36}$$

Note que

$$\sum_{i=1}^{n} (A_{t_i} - A_{t_{i-1}})^2 \leqslant \sup_i |A_{t_i} - A_{t_{i-1}}| \sum_{i=1}^{n} |A_{t_i} - A_{t_{i-1}}| \leqslant \sup_i |A_{t_i} - A_{t_{i-1}}| V_A,$$

onde V_A é a variação de A. Como A é contínuo, A é uniformemente contínuo em $[0, T]$, de forma que o supremo acima vai a 0 qtp e, portanto, vai a 0 em probabilidade. Similarmente,

$$\begin{aligned} \sum_{i=1}^{n} (A_{t_i} - A_{t_{i-1}})(M_{t_i} - M_{t_{i-1}}) &\leqslant \\ \sup_i |M_{t_i} - M_{t_{i-1}}| \sum_{i=1}^{n} |A_{t_i} - A_{t_{i-1}}| &\leqslant \sup_i |M_{t_i} - M_{t_{i-1}}| V_A. \end{aligned}$$

O mesmo raciocínio se aplica, e concluímos que esse processo também vai a 0 em probabilidade. Assim, tomando o limite em probabilidade em (1.36), chegamos a nosso resultado. □

Teorema 1.5.3. *Sejam M e N martingales locais, contínuos e quadrado integráveis em $[0,T]$. Então*

$$\langle M,N\rangle_t = \underset{n\to\infty}{\mathrm{p}\lim}\sum_{i=1}^{n}(M_{t_i}-M_{t_{i-1}})(N_{t_i}-N_{t_{i-1}}), \tag{1.37}$$

onde o limite em probabilidade acima se dá quando percorremos uma sequência qualquer de partições π_n de $[0,t]$ com $|\pi_n| \to 0$

Demonstração. Vamos mostrar (1.37) para o tempo T, sendo os outros casos análogos. Consideramos uma sequência de partições π_n de $[0,T]$, cujos elementos serão genericamente denotados por $0 = t_0 < \ldots < t_n = T$, e tal que $|\pi_n| \to 0$. Também denotaremos M_{t_i} por M_i, analogamente para N, por simplicidade. Usando a equação (1.28) e o Teorema 1.5.1, temos

$$\begin{aligned}
&\langle M,N\rangle_T = \frac{1}{2}(\langle M+N\rangle - \langle M\rangle - \langle N\rangle) = \\
&\frac{1}{2}\underset{n\to\infty}{\mathrm{p}\lim}\sum_{i=1}^{n}\Big(((M_i+N_i)-(M_{i-1}+N_{i-1}))^2-(M_i-M_{i-1})^2-(N_i-N_{i-1})^2\Big).
\end{aligned} \tag{1.38}$$

Abrindo a expressão dentro do somatório,

$$\begin{aligned}
&((M_i+N_i)-(M_{i-1}+N_{i-1}))^2-(M_i-M_{i-1})^2-(N_i-N_{i-1})^2 = \\
&(M_i+N_i)^2+(M_{i-1}+N_{i-1})^2-2(M_i+N_i)(M_{i-1}+N_{i-1}) \\
&-M_i^2-M_{i-1}^2+2M_iM_{i-1}-N_i^2-N_{i-1}^2+2N_iN_{i-1} = \\
&M_i^2+N_i^2+2M_iN_i+M_{i-1}^2+N_{i-1}^2+2M_{i-1}N_{i-1}-2M_iM_{i-1}-2M_iN_{i-1} \\
&-2N_iM_{i-1}-2N_iN_{i-1}-M_i^2-M_{i-1}^2+2M_iM_{i-1}-N_i^2-N_{i-1}^2+2N_iN_{i-1} \\
&=2(M_iN_i+M_{i-1}N_{i-1}-M_iN_{i-1}-M_{i-1}N_i) = 2(M_i-M_{i-1})(N_i-N_{i-1}).
\end{aligned} \tag{1.39}$$

Substituindo agora (1.39) em (1.38), temos o resultado. □

Corolário 1.5.4. *Sejam X, Y semimartingales contínuos. Então,*

$$\langle X,Y\rangle_t = \underset{n\to\infty}{\mathrm{p}\lim}\sum_{i=1}^{n}(X_{t_i}-X_{t_{i-1}})(Y_{t_i}-Y_{t_{i-1}}). \tag{1.40}$$

Demonstração. Análoga à prova de 1.5.2. □

Temos também duas propriedades diretas da variação cruzada:

Proposição 1.5.5. *Sejam X, Y e Z semimartingales contínuos, e $c \in \mathrm{R}$. Então:*

1. $\langle X+cY,Z\rangle_t = \langle X,Z\rangle_t + c\langle Y,Z\rangle_t$.

2. $\langle X,Y\rangle_t = \langle Y,X\rangle_t$.

Demonstração. Ambas seguem diretamente do Teorema (1.5.3). □

Proposição 1.5.6. *Seja M_t^n uma sequência de martingales quadrado integráveis em $[0,T]$ e M_t um martingale quadrado integrável (todos iniciado em 0), tais que M_T^n converge, em L^2, para M_T. Então, $\langle M^n\rangle_T$ converge a $\langle M\rangle_T$ em L^1.*

Demonstração. De fato, pela definição da variação quadrática, temos que

$$(M_t^n - M_t)^2 - \langle M^n - M\rangle_t$$

é um martingale. Com isso, as esperanças de cada um desses processos é a mesma. O resultado sobre a convergência segue claramente. □

Corolário 1.5.7. *Seja M_t^n uma sequência de martingales quadrado integráveis em $[0,T]$, M_t e N_t martingales quadrado integráveis. Então, se $M_T^n \to M_T$ em L^2, então $\langle M^n, N\rangle_T \to \langle M, N\rangle_T$ em L^1.*

Demonstração. Só usar a forma polar da variação quadrática conjunta. □

Também temos a seguinte desigualdade

Proposição 1.5.8 (Desigualdade de Kunita-Watanabe)**.** *Sejam X_t e Y_t martingales locais, e sejam f e g processos previsíveis tais que $\int_0^T |f_s|^2 d\langle X\rangle_t < \infty$ e $\int_0^T |g_s|^2 d\langle Y\rangle_t < \infty$. Então*

$$|\int_0^T f_t g_t d\langle X, Y\rangle_t| \leqslant (\int_0^T |f_t|^2 d\langle X\rangle_t)^{1/2} (\int_0^T |g_t|^2 d\langle Y\rangle_t)^{1/2}$$

Demonstração. Ver (CATUOGNO, 2013, Teorema 34) □

Voltando, para nossos processos de Itô $dX_t = a_t dt + b_t^1 dW_t^1 + \ldots + b^n dW_t^n$, precisamos, pela definição da variação quadrática de semimartigales e da proposição 1.5.5, aprender a calcular apenas variações do tipo

$$\langle \int_0^t f(s) dW_s^1, \int_0^t g(s) dW_s^2 \rangle$$

E isso que veremos agora:

Teorema 1.5.9. *Considere os martingales locais dados por $\int_0^t f(s)dW_s^1$ e $\int_0^t g(s)dW_s^2$. Então*

$$\langle \int_0^\cdot f(s) dW_s^1, \int_0^\cdot g(s) dW_s^2 \rangle_t = \int_0^t f(s) g(s) d\langle W^1, W^2\rangle_s. \tag{1.41}$$

Antes da demonstração, precisamos de um lema auxiliar

Lema 1.5.10. *Considere o martingale local* $\int_0^t f(s)dW_s$ *e o martingale quadrado integrável* N*. Então*

$$\langle \int_0^\cdot f(s)dW_s^1, N\rangle_t = \int_0^t f(s)d\langle W,N\rangle_s. \tag{1.42}$$

Demonstração. Vamos mostrar que os processos são iguais no tempo T, sendo análogo para os outros. Comecemos supondo que $f \in H^2$ além de ser simples, ou seja, do tipo

$$f(\omega,t) = \sum_{i=0}^{n-1} \xi_i(\omega)1_{(t_i,t_{i+1}]}(t).$$

Para essa partição, note que

$$\sum_{i=0}^{n-1}(\int_{t_i}^{t_{i+1}} f(s)W_s)(N_{t_{i+1}} - N_{t_i}) = \sum_{i=0}^{n-1} \xi_i(W_{t_{i+1}} - W_{t_i})(N_{t_{i+1}} - N_{t_i}) =$$
$$\int_0^T f(s)d\langle W,N\rangle_s,$$

pela definição da integral de Lebesgue-Stieltjes em relação à função de variação limitada $\langle W,N\rangle_t$. Note que a expressão acima se mantém a mesma se tomarmos partições mais refinadas que a utilizada. Dessa forma, tomando o limite sobre todas essas partições,

$$\langle \int_0^\cdot f(s)dW_s, N\rangle_T = \int_0^T f(s)d\langle W,N\rangle_s. \tag{1.43}$$

Agora, consideramos $f \in H^2$ não simples. Então, há uma sequência de $f^n \in H^2$ de processos simples tal que $\int_0^T f^n dW_s \to \int_0^T f dW_s$ em L^2. Neste caso, se substituirmos cada f_n em (1.43), temos que o primeiro membro dessa equação tende em L_1 a $\langle \int_0^\cdot f(s)dW_s, N\rangle_T$ pelo corolário 1.5.7. Agora, estudemos a convergência do segundo membro. Pela desigualdade de Kunita-Watanabe,

$$|\int_0^T (f(s) - f^n(s))d\langle W,N\rangle_s| \leqslant (\int_0^T (f_n - f)^2 d\langle W\rangle_t)^{1/2}(\int_0^T 1^2 d\langle N\rangle_t)^{1/2}$$
$$= (\int_0^T (f_n - f)^2 dt)^{1/2}\langle N\rangle_T^{1/2},$$

onde usamos antecipadamente, mas sem prejuízos do tipo argumentação circular, o teorema 1.5.12. Tomando esperanças, pela desigualdade de Cauchy-Schwarz,

$$\mathrm{E}|\int_0^T (f(s) - f^n(s))d\langle W,N\rangle_s| \leqslant \mathrm{E}[\int_0^T (f_n - f)^2 dt]^{1/2}\mathrm{E}[\langle N\rangle_T]^{1/2}.$$

A última esperança é finita e a penúltima vai a 0, pela escolha dos f_n. Com isso, mostramos que quando avaliamos (1.43) em f_n, seus dois membros tendem em L^1 à mesma expressão

(1.43), mas agora avaliada em f. Demonstramos então o teorema para qualquer $f \in H^2$. Falta considerarmos o caso em que $f \in H^2_{loc}$. Para tal f, temos

$$\mathrm{P}\{\omega \mid \int_0^T |f(\omega,t)|^2 dt < \infty\} = 1, \tag{1.44}$$

e temos uma sequência f_n em H^2 tal que $f_n \to f$ para todos (ω, t). Ainda, levando em conta que f_n são do tipo $f_n(\omega,t) = f(\omega,t)1_{\{\tau_n(\omega)\geqslant t\}}$, é fácil ver que $|f_n(\omega,t) - f(\omega,t)|^2 \leqslant |f(\omega,t)|^2$, e esse segundo processo é integrável em $[0,T]$ para quase todo ω. Concluímos então, pelo teorema da convergência dominada, que para quase todo ω as funções $f_n(\omega)$ tendem a $f(\omega)$ em $L^2[0,T]$. Com isso, temos a seguinte convergência ω-qtp:

$$\Big(\int_0^T (f_n(\omega,t) - f(\omega,t))^2 dt\Big)^{1/2} \to 0. \tag{1.45}$$

Como estamos num espaço de medida (probabilidade) finita, a convergência (1.45) se dá também em probabilidade. Lembremos então do seguinte resultado:

Claim: Se $X_n, Y_n \to X, Y$ em probabilidade, respectivamente, e $f : \mathrm{R}^2 \to R$ é contínua, então $f(X_n, Y_n) \to f(X,Y)$ em probabilidade.

A afirmação acima e a observação após (1.45) nos faz concluir então que $\big(\int_0^T (f_n(\omega,t) - f(\omega,t))^2 dt\big)^{1/2} \langle N \rangle_T^{1/2} \to 0$ em probabilidade. Com isso, pela desigualdade de Kunita-Watanabe, $\int_0^T (f_n(s) - f(s)) d\langle W, N\rangle_s \to 0$ também em probabilidade. Este foi o fato 1. Agora, a integral de Itô $\int_0^T f dW$ é definida como (a versão contínua) do limite qtp de $\int_0^T f_n dW$. Essa convergência também ocorre da seguinte maneira: os processos parados $(I(f_n)_t^{\tau_m})$ convergem em H^2 ao processo parado $(I(f)_t^{\tau_m})$, todos martingales. Esse é o fato 2. Por isso, concluímos que $\langle I(f_n)_t, N\rangle \to \langle I(f)_t, N\rangle$ em probabilidade (ver (CATUOGNO, 2013, Teorema 29.2.a)). Assim, substituindo f_n na equação (1.43) e tomando limites em probabilidade dos dois lados, pelos fatos 1 e 2 nós obtemos a mesma expressão (1.43) avaliada em f, e demonstramos o resultado deste lema. □

Voltemos à demonstração do Teorema 1.5.9:

Demonstração. A demonstração segue facilmente de duas aplicações do lema anterior. Comece considerando o caso em que $g \in H^2$, ou seja, $\int_0^t g(s) dW_s^2$ é martingale quadrado integrável:

$$\begin{aligned}\langle \int_0^\cdot f(s) dW_s^1, \int_0^\cdot g(s) dW_s^2 \rangle_t = \int_0^t f(s) d\langle W^1, \int_0^\cdot g(u) W_u^2 \rangle_s = \\ \int_0^t f(s) g(s) d\langle W^1, W^2\rangle_s.\end{aligned} \tag{1.46}$$

Agora, se $g \in H^2_{loc}$, tome uma sequência τ_n de tempos de parada que localiza g para $g_n \in H^2$. Então a igualdade em (1.46) vale para todo g_n. Tomando limites qtp (passando por uma subsequência se necessário), obtemos o resultado desejado. □

Agora, resumimos nossa problema de encontrar os integradores presentes na fórmula de Itô multidimensional a encontrar a variação conjunta entre movimentos Brownianos.

Teorema 1.5.11. *Sejam X e Y movimentos brownianos independentes em $[0, T]$. Então*

$$\langle X, Y \rangle_t = 0, \qquad \forall t \in [0, T]$$

Demonstração. Seja $\Delta = \{t_0, \ldots, t_n$ uma partição de $[0, T]$. Vamos denotar por X_i a variável aleatória X_{t_i}, por Y_i a variável aleatória Y_{t_i} e por $|\Delta|$ o tamanho da partição. Calculemos

$$\mathrm{E}[(\sum_{i=1}^{n}(X_i - X_{i-1})(Y_i - Y_{i-1}))^2].$$

Note que temos termos de 2 tipos na soma:

$$\begin{aligned}&(X_i - X_{i-1})^2(Y_i - Y_{i-1})^2, \qquad e\\ &(X_i - X_{i-1})(Y_i - Y_{i-1})(X_j - X_{j-1})(Y_j - Y_{j-1})), \qquad i \neq j,\end{aligned}$$

os quais chamaremos de U_i e $U_{i,j}$, respectivamente. Assim,

$$\mathrm{E}[(\sum_{i=1}^{n}(X_i - X_{i-1})(Y_i - Y_{i-1}))^2] = \sum_{i \neq j} \mathrm{E}[U_{i,j}] + \sum_{i=1}^{n} \mathrm{E}[U_i].$$

Por independência de X e Y e dos seus respectivos incrementos, temos

$$\mathrm{E}[U_{i,j}] = \mathrm{E}(X_i - X_{i-1})\mathrm{E}(X_j - X_{j-1})\mathrm{E}(Y_i - Y_{i-1})\mathrm{E}(Y_j - Y_{j-1}) = 0 \cdot 0 \cdot 0 \cdot 0 = 0.$$

Além disso, por independência de X e Y,

$$\mathrm{E}[U_i] = \mathrm{E}[(X_i - X_{i-1})^2]\mathrm{E}[(Y_i - Y_{i-1})^2] = (t_i - t_{i-1})^2.$$

Dessa forma,

$$\mathrm{E}[(\sum_{i=1}^{n}(X_{i-1} - X_i)(Y_{i-1} - Y_i))^2] = \sum_{i=1}^{n}(t_i - t_{i-1})^2 \leqslant |\Delta| \sum_{i=1}^{n}(t_i - t_{i-1}) = |\Delta| T.$$

Considerando uma sequência de partições tal que $|\Delta_n| \to 0$, concluímos que

$$\sum_{i=1}^{n}(X_i - X_{i-1})(Y_i - Y_{i-1}) \to 0, \text{quando}\, n \to \infty, \text{ em} L^2(\mathrm{P}).$$

Assim,

$$\sum_{i=1}^{n}(X_i - X_{i-1})(Y_i - Y_{i-1}) \to 0, \text{quando}\, n \to \infty, \text{em probabilidade},$$

e obtemos que $\langle X, Y \rangle_T = 0$. De forma análoga, $\langle X, Y \rangle_t = 0$, e o resultado segue. □

Teorema 1.5.12. *Seja W_t um movimento Browniano em $[0,T]$. Então, $\langle W\rangle_t = t$.*

Demonstração. Poderíamos usar o teorema 1.5.1, porém há uma forma mais prática. Considere o submartingale W_t^2. Lembremos que, para $s < t$, temos, pelo teorema A.3.6, que

$$\begin{aligned} \mathrm{E}[W_t^2 - t \mid \mathcal{F}_s] &= \mathrm{E}[(W_t^2 - W_s^2) \mid \mathcal{F}_s] - t + W_s^2 \\ = \mathrm{E}[(W_t - W_s)^2 \mid \mathcal{F}_s] - t + W_s^2 &= W_s^2 - t + (t-s) = W_s^2 - s. \end{aligned}$$

Assim, $W_t^2 - t$ é um martingale e, pela unicidade dada na decomposição de Doob-Meyer, $\langle W\rangle_t = t$, como anunciado. □

Exemplo 1.5.13. *Calculemos a integral simples $\int_0^t W_t dW_t$ e percebamos como o cálculo de Itô difere do cálculo usual. Considere a função $f(x) = x^2$. Então, pela formula de Itô,*

$$\begin{aligned} &d(f(W_t)) = f'(W_t)dW_t + \frac{1}{2}f''(W_t)d\langle W\rangle_t = 2W_t dW_t + \frac{1}{2}2dt \\ &\Rightarrow \int_0^T W_t dW_t = \frac{1}{2}(W_T^2 - T). \end{aligned}$$

Em resumo, podemos calcular variações quadráticas de processos de Itô da seguinte forma: dados $dX = adt + b^1 dW_t^1 + \ldots + b^n dW_t^n$ e $dY_t = \tilde{a}dt + \tilde{b}^1 d\tilde{W}_t^1 + \ldots + \tilde{b}^m d\tilde{W}_t^m$, então $d\langle X, Y\rangle = dX \cdot dY$, onde $\cdot$ deve ser encarada como uma multiplicação distributiva usual e as regras de multiplicação seguirem a convenção $dt \cdot dt = 0$, $dW_t^i \cdot dW_t^j = \lambda_{i,j}dt$, onde $\lambda_{i,j} = 0$ se W_t^i e W_t^j forem movimentos Brownianos independentes e $\lambda_{i,j} = 1$ se forem o mesmo movimento Browniano.

1.6 Equações Diferenciais Estocásticas

De posse de uma nova teoria de integração, temos uma nova teoria de equações diferenciais. Uma SDE (stochastic differential equation) é uma equação e uma condição inicial do tipo

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \qquad X_0 = x.$$

Como não poderia deixar de ser, uma solução para a SDE acima em $[0,T]$ é um processo adaptado $\Omega \times [0,T]$ tal que $X_t = x + \int_0^t b(s, X_s)ds + \sigma(s, X_s)dW_s$, para todo $t \in [0,T]$. Enunciaremos o resultado padrão de existência e unicidade para soluções de SDEs sujeitas a determinadas restrições, mostraremos algumas de suas propriedades e resolveremos um tipo especial de SDE muito útil, chamada linear, através da fórmula de Itô,

Teorema 1.6.1 (Existência e Unicidade de Solução de SDEs)**.** *Sejam* $b : [0,T] \times \mathrm{R}^d \to \mathrm{R}^d$ *e* $B : [0,T] \times \mathrm{R}^d \to R^{d\times n}$ *funções contínuas que satisfazem a condição de Lispchitz uniforme em* t *e uma limitação de crescimento:*

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leqslant L|x-y|; \tag{1.47}$$

$$|b(t,x)| + |\sigma(t,x)| \leqslant L(1+|x|), \tag{1.48}$$

para alguma constante $L > 0$*. Então, existe um único processo em* L^2 *que seja solução da SDE*

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dW_t, \qquad X_0 = x. \tag{1.49}$$

Demonstração. (EVANS, 2012, Seção 5.2) □

Teorema 1.6.2. *Seja* $X_t, t \in [0,T]$ *a solução da SDE* (1.49)*, prevista no teorema 1.6.1. Então* X_t *satisfaz a propriedade de Markov, isto é, para qualquer* $g : \mathrm{R}^d \to \mathrm{R}$ *limitada e Borel-mensurável, temos*

$$\mathrm{E}[g(X_{t+h})|\mathcal{F}_t] = \mathrm{E}_{h,X_h}(g(X_t)).$$

Demonstração. (KUO, 2005, Teorema 10.6.1) □

Teorema 1.6.3 (Majoramento para uma solução.)**.** *Sob as hipóteses do teorema 1.6.1, a solução providenciada por este teorema satisfaz a seguinte condição de integrabilidade*

$$\mathrm{E}[\sup_{t\in[0,T]} |X_t|^p] \leqslant C_T(1+|x|^p),$$

onde C_T *é uma constante e estamos considerando* $p > 1$*.*

Exemplo 1.6.4 (Movimento Browniano Geométrico)**.** *Considere a SDE*

$$dX_t = X_t(\mu dt + \sigma dW_t), X_0 = a. \tag{1.50}$$

Agora, considere o processo de Itô,

$$X_t = a\exp\{(\mu - \frac{\sigma^2}{2})t + \sigma W_t\}. \tag{1.51}$$

Podemos escrevê-lo como $f((\mu - \frac{\sigma^2}{2})t + \sigma W_t) = f(Y_t)$*, onde* $f(x) = a\exp(x)$ *e* $Y_t = (\mu - \frac{\sigma^2}{2})t + \sigma W_t$*. Então, pela fórmula de Itô:*

$$dX_t = f'(Y_t)dY_t + \frac{1}{2}f''(Y_t)d\langle Y\rangle_t = X_t((\mu - \frac{\sigma^2}{2})dt + \sigma dW_t) + \frac{1}{2}X_t\sigma^2 dt =$$
$$X_t(\mu dt + \sigma dW_t).$$

Ainda, $X(0) = a$*, e, portanto, o processo dado em* (1.51)*, chamado de* ***movimento Browniano geométrico****, resolve a SDE linear* (1.50)*. Veremos depois que essa equação modela o preço de uma ação no mercado, e por isso a sua importância para nós.*

2 Backward Stochastic Differential Equations

Desenvolveremos neste capítulo um tipo novo de equação estocástica, que chamaremos de BSDEs. Nos preocuparemos com duas vertentes de aplicações dessas equações: a representação de EDPs semilineares; e a sua utilização para solução do problema de se precificar derivativos num mercado de valores, o que é um prelúdio para a imensidade de aplicações financeiras que as BSDEs possuem e cuja nossa referência padrão é (KAROUI; PENG; QUENEZ, 1997).

2.1 Representação de Feynman-Kac

Desenvolveremos aqui a representação (estocástica) de Feynman-Kac para soluções de EDPs parabólicas lineares, de forma heurística e depois rigorosamente. Veremos ao se tentar generalizar essa representação para EDPs parabólicas semilineares, somos conduzidos a resolver um tipo de equação diferencial estocástica, onde a solução procurada é um par de processos, e nos é fornecida uma condição terminal, aspecto fundamental. A representação de equações semilineares, apesar de não ser o foco do capítulo ou mesmo do trabalho, proporciona um bom exemplo de surgimento espontâneo das BSDEs. A ideia sobre como tentar desenvolver uma representação para EDPs, começando por uma função definida em termos de uma esperança e vendo que tipo de equação diferencial determinística ela satisfaz, segue da primeira parte do artigo de referência (PERKOWSKI, 2011).

Seja X_t o processo de Itô em R^d dado por

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t, \tag{2.1}$$

onde W_t é um movimento Browniano padrão n-dimensional (e portanto $b : R^n \to R^n$, $\sigma : R^n \to R^{d\times n}$). Estamos supondo b e σ satisfazendo as condições do Teorema 1.6.1, de forma que a SDE (2.1) tem uma única solução bem definida (e em L^2). Escreveremos $X_t = (X_t^1, \ldots, X_t^d)$. Chamamos de **gerador infinitesimal do processo** X_t o operador diferencial $\mathcal{L}^X = \mathcal{L} : C^2(\mathrm{R}^d) \to C(\mathrm{R})$ dado por

$$(\mathcal{L}v)(x) := \sum_{i=1}^{d} b_i(x)\partial_i v(x) + \frac{1}{2}\sum_{i,j=1}^{d} (\sigma\sigma^T)_{ij}(x)\partial_{ij}v(x). \tag{2.2}$$

Destacamos tal operador pois, para qualquer função $v \in C^2(\mathrm{R}^d)$, temos

$$dv(X_t) = \mathcal{L}v(X_t) + \sum_{i=1}^{d}\sum_{j=1}^{n} \partial_i v(X_t)\sigma_{ij}(X_t)dW_j = \mathcal{L}v(X_t) + \nabla v(X_t)\sigma(X_t)dW_t. \tag{2.3}$$

De fato, levando em conta que

$$d\langle X^l, X^m\rangle_t = (b_l dt + \sum_{k=1}^{n} \sigma_{lk} dW_k) \cdot (b_m dt + \sum_{q=1}^{n} \sigma_{mq} dW_q)) =$$
$$\sum_{k,q=1}^{n} \sigma_{lk}\sigma_{mq}(dW_k \cdot dW_q) = \sum_{k=1}^{n} \sigma_{lk}\sigma_{mk} dt = (\sigma\sigma^T)_{lm} dt,$$

temos, pela fórmula de Itô, que

$$dv(X_t) = dv(X_t^1, \ldots, X_t^d) = \sum_{k=1}^{d} \partial_k v(X_t) dX_t^k + \frac{1}{2}\sum_{l,m=1}^{d} \partial_{lm} v(X_t) d\langle X^l, X^m\rangle_t =$$
$$\sum_{k=1}^{d} \partial_k v(X_t)(b_k(X_t)dt + \sum_{l=1}^{n} \sigma_{kl}(X_t) dW_t^l) + \frac{1}{2}\sum_{l,m=1}^{d} \partial_{lm} v(X_t)(\sigma\sigma^T)_{lm}(X_t)dt =$$
$$\mathcal{L}v(X_t)dt + \sum_{i=1}^{d}\sum_{j=1}^{n} \partial_i v(X_t)\sigma_{ij}(X_t) dW_t^j = \mathcal{L}v(X_t)dt + \nabla v(X_t)\sigma(X_t)dW_t.$$

De forma análoga, se considerarmos processos de Itô X_t que satisfazem SDEs de coeficientes mais gerais (que também chamaremos de b e σ),

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \tag{2.4}$$

onde, novamente, b e σ satisfazem as hipóteses do Teorema 1.6.1, podemos definir a família de operadores diferenciais $\mathcal{L}_t^X = \mathcal{L}_t : C^2(\mathrm{R}^d) \to C(\mathrm{R}^d)$, $t \in [0, T]$, por

$$\mathcal{L}_t v(x) = \sum_{i=1}^{d} b_i(t, x)\partial_i v(x) + \frac{1}{2}\sum_{i,j=1}^{d} (\sigma\sigma^T)_{ij}(t, x)\partial_{ij} v(x).$$

Neste caso, se $v \in C^{1,2}([0, T] \times \mathrm{R}^d)$, então

$$dv(t, X_t) = (\partial_t v(t, X_t) + \mathcal{L}v(t, X_t))dt + \nabla_x v(t, X_t)\sigma(t, X_t)dW_t. \tag{2.5}$$

Seja $\psi \in C^2(\mathrm{R}^d)$ e defina $u(t, x) := \mathrm{E}_x(\psi(X_t))$, X_t dado por (2.1), onde o índice x significa que estamos considerando a medida de probabilidade associada ao movimento Browniano iniciado em x. Supondo que u seja suficientemente “bem comportada”, vamos encontrar uma EDP satisfeita por u, heuristicamente. Notemos que

$$u(t + h, x) = \mathrm{E}_x(\psi(X_{t+h})) = \mathrm{E}_x(E_x[\psi(X_{t+h})|\mathcal{F}_h]). \tag{2.6}$$

Usando o fato de que X satisfaz a propriedade de Markov (Teorema 1.6.2), a definição de u e (2.3), temos

$$\mathrm{E}_x(\mathrm{E}_x[\psi(X_{t+h})|\mathcal{F}_h]) = \mathrm{E}_x(\mathrm{E}_{X_h}(\psi(X_t))) = \mathrm{E}_x(u(t, X_h)) =$$
$$\mathrm{E}_x(u(t, X_0) + \int_0^h \mathcal{L}u(t, X_s)ds + \sum_{i=1}^{d}\sum_{j=1}^{n}\int_0^h \partial_i u(t, X_s)\sigma_{ij}(X_s)dW_s^j) =$$
$$u(t, x) + \mathrm{E}_x(\int_0^h \mathcal{L}u(t, X_s)ds).$$

Dessa forma,

$$\frac{u(t+h,x)-u(t,x)}{h} = \mathrm{E}_x(\frac{1}{h}\int_0^h \mathcal{L}u(X_s)ds).$$

Tomando o limite quando $h \to 0$, supondo que possamos trocar o limite com a esperança e usando o teorema do valor médio para integrais,

$$\begin{aligned} u_t(t,x) &= \mathcal{L}u(t,x) \\ u(0,x) &= \psi(x). \end{aligned} \tag{2.7}$$

Assim, somos direcionados a entender que a função $\mathrm{E}_x(\psi(X_t))$ satisfaz o problema de Cauchy acima (2.7). Re-escrevendo u, temos $d/dt\mathrm{E}_x(\psi(X_t)) = \mathcal{L}\mathrm{E}_x(\psi(X_t))$, chamada de **Equação backward de Kolmogorov**. Ela nos leva também a uma representação de um problema de contorno:

$$\begin{aligned} \mathcal{L}v(x) &= 0, v \in D, \\ v(x) &= f(x), x \in \partial D, \end{aligned} \tag{2.8}$$

onde D é um aberto limitado de R^d. Seja τ_D o tempo de parada

$$\tau_D = \inf\{t \geqslant 0 | X_t \notin D\},$$

e considere a função $v : \bar{D} \to R$, $x \mapsto \mathrm{E}_x(f(X_{\tau_D}))$. Note que

1. $\mathcal{L}v = \mathcal{L}\mathrm{E}_x(f(X_{\tau_D})) = d/dt\mathrm{E}_x(f(X_{\tau_D})) = 0$, pela equação backward, $\forall x \in D$;
2. para todo $x \in \partial D$, $v(x) = \mathrm{E}_x(f(X_{\tau_D})) = \mathrm{E}_x(f(X_0)) = \mathrm{E}_x(f(x)) = f(x)$, já que $X_{\tau_D} \in \partial D$.

Assim, vemos que $v(x) = \mathrm{E}_x(f(X_{\tau_D}))$ resolve o problema de contorno (2.8).

Obtemos então, a princípio, soluções para as equações (2.7) e (2.8) dadas através de esperanças de processos de Itô. Uma das vantagens desse tipo de representação de soluções de EDP, num campo teórico, é o corolário direto sobre unicidade de soluções, que veremos com detalhes abaixo. Por outro lado, no campo numérico, tais representações fornecem, em tese, uma outra forma de abordagem de aproximação e simulação para essas soluções, como métodos numéricos probabilísticos, por exemplo Monte-Carlo. Ver, em particular, (Zhou; Cai, 2015) para mais detalhes.

Em matemática, usualmente introduzimos um resultado de forma ingênua e suas possíveis aplicações, para depois demonstrá-lo de maneira rigorosa, apresentar exemplos e tentar generalizá-lo. Seguiremos agora essa sequência, com dois teoremas, conhecidos como *fórmulas de Feynman-Kac*, que providenciam representações estocásticas de soluções de EDPs, dadas por problemas de Cauchy ou Dirichlet, satisfeitas determinadas condições. Essa será a parte de tornar rigorosa a intuição obtida acima. Os seguintes enunciados e demonstrações seguem de perto o tratamento dado em (KARATZAS; SHREVE, 1988, seção 5.7), com algumas modificações a fim de detalhar mais as contas mais abertas.

Teorema 2.1.1 (Representação de Feynman-Kac)**.** *Sejam* $f : \mathrm{R}^d \to \mathrm{R}$*,* $g : [0,T] \times \mathrm{R}^d \to \mathrm{R}$*, e* $k : [0,T] \times \mathrm{R}^d \to [0,\infty)$ *funções contínuas satisfazendo:*

(i) $|f(x)| \leqslant L(1 + |x|^{2\lambda})$ *ou* $f(x) \geqslant 0$*,* $\forall x \in \mathrm{R}^d$*,*

(ii) $|g(t,x)| \leqslant L(1 + |x|^{2\lambda})$ *ou* $g(t,x) \geqslant 0$*,* $\forall t \in [0,T], x \in \mathrm{R}^d$*,*

onde $L > 0$*,* $\lambda \geqslant 1$ *são constantes. Suponha também que* $v : [0,T] \times \mathrm{R}^d \to \mathrm{R}$ *seja contínua, de classe* $C^{1,2}$ *em* $[0,T) \times \mathrm{R}^d$*, e satisfaça o problema de Cauchy*

$$\begin{aligned} &\mathcal{L}_t v + \frac{dv}{dt} - kv = -g, \forall (t,x) \in [0,T) \times \mathrm{R}^d, \\ &v(T,x) = f(x), \forall x \in \mathrm{R}^d, \end{aligned} \tag{2.9}$$

além da condição de crescimento

$$\sup_{t \in [0,T]} |v(t,x)| \leqslant M(1 + |x|^{2\mu}); x \in \mathrm{R}^d, \tag{2.10}$$

onde $M > 0$ *e* $\mu \geqslant 1$ *são constantes. Então* $v(t,x)$ *admite a representação estocástica*

$$v(t,x) = \mathrm{E}_{t,x}\Big[f(X_T) \exp\Big\{ -\int_t^T k(u,X_u)du\Big\} + \int_t^T g(s,X_s) \exp\Big\{ -\int_t^s k(u,X_u)du\Big\}ds\Big], \tag{2.11}$$

$$\forall (t,x) \in [0,T] \times \mathrm{R}^d.$$

Demonstração. Seja Y_s o processo $\exp\big\{ -\int_t^s k(u,X_u)du\big\}$, e vamos aplicar a fórmula de Itô para $Z_s = v(s,X_s)Y_s$. Temos que

$$dZ_s = d(v(X_s,s))Y_s + v(X_s,s)dY_s + d\langle v(\cdot,X_\cdot),Y\rangle_s.$$

Note que

$$d(-\int_t^s k(u,X_u)du) = -k(s,X_s)ds,$$

de forma que

$$\begin{aligned} &dY_s = -\exp\Big\{ -\int_t^s k(u,X_u)du\Big\}k(s,X_s)ds + \\ &\frac{1}{2}\exp\Big\{ -\int_t^s k(u,X_u)du\Big\}(k(s,X_s)ds \cdot k(s,X_s)ds) = \\ &-\exp\Big\{ -\int_t^s k(u,X_u)du\Big\}k(s,X_s)ds = -Y_s k(s,X_s)ds. \end{aligned}$$

Assim,

$$d\langle v(\cdot,X_\cdot),Y\rangle_s = d(v(X_s,s)) \cdot dY_s = 0,$$

pois esse último termo apresenta apenas a "diferencial" ds. Por (2.5), e utilizando a equação satisfeita por v,

$$d(v(s,X_s)) = (\frac{d}{dt} + \mathcal{L}_t)v(s,X_s)ds + \nabla_x v(s,X_s)\sigma(s,X_s)dW_s =$$
$$(k(s,X_s)v(s,X_s) - g(s,X_s))ds + \nabla_x v(s,X_s)\sigma(s,X_s)dW_s.$$

Com isso,

$$dZ_s = v(s,X_s)dY_s + Y_s dv(s,X_s) =$$
$$-v(s,X_s)Y_s k(s,X_s)ds + Y_s(k(s,X_s)v(s,X_s) - g(s,X_s))ds +$$
$$Y_s\nabla_x v(s,X_s)\sigma(s,X_s)dW_s = -Y_s g(s,X_s)ds + Y_s\nabla_x v(s,X_s)\sigma(s,X_s)dW_s.$$

Sabendo que $Z_u = Z_t + \int_t^u dZ_s$ para qualquer $u \in [t,T]$, e usando a definição de Z_s, temos

$$v(u,X_u)Y_u = v(t,X_t)Y_t - \int_t^u Y_s g(s,X_s)ds + \int_t^u Y_s\nabla_x v(s,X_s)\sigma(s,X_s)dW_s.$$

Tomando X_t o processo que satisfaz a SDE (2.4) com a condição inicial $X_t = x$, denotando por $\mathrm{E}_{t,x}$ a medida de probabilidade associada ao movimento Browniano que se inicia em x no tempo t, notando que $Y(\omega,t) = 1$ e substituindo u pelo tempo de parada $T \wedge \tau_n$, onde

$$\tau_n = \inf\{s \geqslant t; |X_s| \geqslant n\},$$

temos

$$v(t,x) = v(T\wedge\tau_n, X_{T\wedge\tau_n})Y_{T\wedge\tau_n} + \int_t^{T\wedge\tau_n} Y_s g(s,X_s)ds$$
$$-\int_t^{T\wedge\tau_n} Y_s\nabla_x v(s,X_s)\sigma(s,X_s)dW_s.$$

Vamos cacular $\mathrm{E}^{t,x}[\int_t^{T\wedge\tau_n} Y_s\nabla_x v(s,X_s)\sigma(s,X_s)dW_s]$. Note que podemos escrever a integral como

$$\int_t^T 1_{\{\tau_n\geqslant s\}}Y_s\nabla_x v(s,X_s),\sigma(s,X_s)dW_s.$$

Vamos considerar o caso $d = 1$ para facilitar os cálculos, sendo o caso geral completamente análogo. Como, por hipótese, v é de classe $C^{1,2}$, existe uma constante C_n tal que

$$|\partial_x v(s,x)| \leqslant C_n, \forall (s,x) \in [t,T]\times B[0,n].$$

Logo,

$$|1_{\{\tau_n\geqslant s\}}Y_s\partial_x v(s,X_s)\sigma(s,X_s)|^2 \leqslant |1_{\{\tau_n\geqslant s\}}\partial_x v(s,X_s)\sigma(s,X_s)|^2$$
$$\leqslant C_n^2|1_{\{\tau_n\geqslant s\}}\sigma(s,X_s)|^2 \leqslant C_n^2K^21_{\{\tau_n\geqslant s\}}(2+2|X_s|^2) \leqslant 2C_n^2K^2(1+n^2),$$

onde a primeira desigualdade vem de $|Y_s| \leqslant 1$, já que k é uma função não negativa; a segunda desigualdade vem da observação sobre a limitação de $\partial_x v$ e da definição de τ_n; e a última da hipótese de limitação sobre σ e do fato de que $(a+b)^2 \leqslant 2a^2 + 2b^2$. Dessa forma, o integrando acima pertence ao espaço $L^2(\Omega \times [t,T])$, de maneira que a integral estocástica considerada é um martingale. Com isso, concluimos que sua esperança é nula. Sendo $v(t,x)$ uma quantidade determinística,

$$
\begin{aligned}
(*) \qquad v(t,x) &= \mathrm{E}_{t,x}\Big[\int_t^{T\wedge\tau_n} g(s,X_s)Y_s ds\Big] \\
&+ \mathrm{E}_{t,x}\Big[v(\tau_n, X_{\tau_n})Y_{\tau_n}1_{\{\tau_n \leqslant T\}}\Big] + \mathrm{E}_{t,x}\Big[f(X_T)Y_T 1_{\{\tau_n > T\}}\Big],
\end{aligned}
$$

onde usamos a condição inicial satisfeita por v. Queremos cacular os limites de cada uma das três eperanças acima quando $n \to \infty$. Antes, notemos que

$$\tau_n \nearrow \infty \,\text{qtp} \Rightarrow T \wedge \tau_n \nearrow T \,\text{qtp},$$

por continidade de X_s. Comecemos com o primeiro termo. Note que se $g(t,x) \geqslant 0$ (por hipótese), então, usando o fato de $Y_s \geqslant 0$, a sequência de variáveis aletórias definidas pela primeira integral é tal que

$$\int_t^{T\wedge\tau_n} g(s,X_s)Y_s ds \nearrow \int_t^T g(s,X_s)Y_s ds.$$

Com isso, podemos usar o teorema da convergência monótona para concluir que

$$\mathrm{E}^{t,x}\Big[\int_t^{T\wedge\tau_n} g(s,X_s)Y_s ds\Big] \nearrow \mathrm{E}^{t,x}\Big[\int_t^T g(s,X_s)Y_s ds\Big].$$

Por outro lado, podemos ter $|g(t,x)| \leqslant L(1+|x|^{2\lambda})$. Neste caso,

$$
\begin{aligned}
&|\int_t^{T\wedge\tau_n} g(s,X_s)Y_s ds| \leqslant \int_t^{T\wedge\tau_n} |g(s,X_s)Y_s| ds \leqslant \int_t^T |g(s,X_s)| ds \leqslant \\
&L\int_t^T 1 + |X_s|^{2\lambda} ds \leqslant L(T-t)(1 + \max_{t\leqslant u\leqslant T} |X_u|^{2\lambda}).
\end{aligned}
$$

Sabemos que

$$\mathrm{E}_{t,x}\Big[\max_{t\leqslant u\leqslant T} |X_u|^{2\lambda}\Big] \leqslant C(1+|x|^{2\lambda})e^{C(T-t)},$$

onde C é uma constante positiva que só depende de λ, K, T, d, de onde concluimos que a sequência $\int_t^{T\wedge\tau_n} g(s,X_s)Y_s ds$ é majorada por uma variável aleatória de esperança finita, e obtemos também, agora pelo teorema da convergência dominada, que

$$\mathrm{E}^{t,x}\Big[\int_t^{T\wedge\tau_n} g(s,X_s)Y_s ds\Big] \to \mathrm{E}^{t,x}\Big[\int_t^T g(s,X_s)Y_s ds\Big].$$

Vamos agora cacular o limite do segundo termo de $(*)$. Perceba que, pela hipótese de limitação em v, (2.10), temos

$$
\begin{aligned}
&|\mathrm{E}_{t,x}[v(\tau_n, X_{\tau_n})1_{\{\tau_n \leqslant T\}}]| \leqslant M\mathrm{E}_{t,x}[(1+|X_{\tau_n}|^{2\mu})1_{\{\tau_n \leqslant T\}}] \\
&\leqslant M(1+n^{2\mu})\mathrm{P}_{t,x}[\tau_n \leqslant T].
\end{aligned}
$$

Caculamos essa última probabilidade pela desigualdade de Chebyshev e pela limitação em $\max|X_s|$:

$$\begin{aligned}
\mathrm{P}_{t,x}[\tau_n \leqslant T] &= \mathrm{P}_{t,x}[\max_{t\leqslant u\leqslant T}|X_u| \geqslant n] \\
&\leqslant n^{-2m}\mathrm{E}_{t,x}[\max_{t\leqslant u\leqslant T}|X_u|^{2m}] \leqslant Cn^{-2m}(1+|x|^{2m})e^{CT},
\end{aligned}$$

onde m pode ser escolhido como quisermos. Com isso,

$$\mathrm{E}^{t,x}|v(\tau_n, X_{\tau_n})1_{\{\tau_n\leqslant T\}}| \leqslant C_1n^{-2m} + C_2n^{2(\mu-m)},$$

e, escolhendo $m > \mu$, essa esperança vai a zero quando $n \to \infty$. Por termos $|Y_s \leqslant 1|$, vemos que o segundo termo em $(*)$ vai a zero. Finalmente, calculemos o limite do terceiro termo em $(*)$. Note que

$$1_{\tau_n > T} \nearrow 1 \,\text{qtp},$$

trivialmente pela continuidade de X_s. Com isso, se tivermos $f(x) \geqslant 0$, então

$$f(X_T)Y_T1_{\tau_n>T} \nearrow f(X_T)Y_T,$$

e

$$\mathrm{E}_{t,x}\Big[f(X_T)Y_T1_{\tau_n>T}\Big] \to \mathrm{E}_{t,x}\Big[f(X_T)Y_T\Big],$$

pelo teorema da convergencia monótona. Por outro lado, se tivermos $|f(t,x)| \leqslant L(1+|x|^{2\lambda})$, então

$$|f(X_T)Y_T1_{\tau_n>T}| \leqslant |f(X_T)| \leqslant L(1+|X_T|^{2\lambda}),$$

e

$$\mathrm{E}_{t,x}[|X_T|^{2\lambda}] \leqslant \mathrm{E}_{t,x}[\max_{t\leqslant u\leqslant T}|X_u|^{2\lambda}] \leqslant C(1+|x|^{2\lambda})e^{C(T-t)} < \infty,$$

onde podemos aplicar agora o teorema da convergência dominada para também concluir a convergência anterior. Dessa forma, podemos tomar o limite quando $n \to \infty$ em $(*)$ e usar a definição de Y_s para concluir finalmente que

$$v(t,x) = \mathrm{E}^{t,x}\Big[f(X_T)\exp\Big\{-\int_t^T k(u,X_u)du\Big\} + \int_t^T g(s,X_s)\exp\Big\{-\int_t^s k(u,X_u)du\Big\}ds\Big]$$

□

Como corolário óbvio, a representação de Feynmam-Kac acima, em particular, nos garante unicidade de solução pro problema de Cauchy sujeita à condição de crescimento $\max_{0\leqslant t\leqslant T}|v(t,x)| \leqslant M(1+|x|^{2\mu}); x \in \mathrm{R}^d$. De fato, quaiquer duas soluções sujeitas à limitação (2.10) serão escritas da mesma forma como (2.11).

Note que a representação acima, apesar de garantir *unicidade*, está sujeita à *existência* da solução. A pergunta natural então é sobre quando a EDP parabólica (2.9) acima possui solução. Estudar os detalhes de quando há solução garantida não é objeto desse trabalho, contudo, por completude, vamos enunciar um resultado padrão. Antes, porém, precisamos estudar um pouco o operador $\mathcal{L}$.

Definição 2.1.2. *Diremos que um operador (linear) $\mathcal{L} : C^2(\mathrm{R}^d) \to C(\mathrm{R}^d)$ é um **operador diferencial de segunda ordem** quando é do tipo*

$$(\mathcal{L}u)(x) = \sum_{i,j=i}^{d} a_{i,j}(x)\partial_{i,j}u(x) + \sum_{i=1}^{d} b_i(x)\partial_i u(x) + c(x)u(x).$$

*Diremos que um operador diferencial de segunda ordem é **elíptico** quando*

$$\sum_{i,j=1}^{d} a_{i,j}(x)v_i v_j > 0, \forall(v_1, \ldots, v_d) \in \mathrm{R}^d - \{0\}. \tag{2.12}$$

Perceba que (2.12) acima pode ser escrita como $v^T(a_{i,i}(x))v > 0$, de forma que, no caso do nosso gerador infinitesimal $\mathcal{L}$, temos a condição equivalente a $v^T(\sigma(x)\sigma^T(x))v = (\sigma^T(x)v)^T(\sigma^T(x)v) = \langle\sigma^T v, \sigma^T v\rangle > 0\ \forall v \neq 0 \Leftrightarrow \sigma(x)$ invertível para todo x. Assim, se σ sempre for invertível (o que requisita, no mínimo, um processo X com a mesma dimensão do movimento Browniano considerado), o operador $\mathcal{L}$ será elíptico. Mais ainda, definindo por $\mathcal{S}(x)$ o espectro da matriz $\sigma(x)\sigma^T(x)$, e levando em conta que $\sigma(x)\sigma(x)^T$ é autoadjuta para todo x, se tivermos que $\inf_x \min(\lambda \in \mathcal{S}(x)) > C > 0$, então teremos que

$$\sum_{i,j=1}^{d} a_{i,j}(x)v_i v_j > L||v||, \qquad L > 0. \tag{2.13}$$

De fato, para cada x, há uma matriz unitária $P(x)$ que diagonaliza $A(x) = \sigma(x)\sigma(x)^T$. Com isso,

$$\begin{aligned}
&\sum_{i,j=1}^{d} a_{i,j}(x)v_i v_j = v^T A(x)v = \\
&v^T\sigma(x)\sigma(x)^T v = v^T P(x)P^T(x)\sigma(x)\sigma(x)^T P(x)P(X)^T v = \\
&(P(x)^T v)^T(P(x)^T\sigma(x)\sigma(x)^T P(x))P(x)^T v = (P(x)^T v)^T D(x)P(x)^T v = \\
&d_1(x)(P(x)^T v)_1^2 + \ldots + d_n(x)(P(x)^T v)_n^2 \geqslant \min\mathcal{S}(x)||P(x)^T v||^2 = \\
&\min\mathcal{S}(x)||v||^2 \geqslant C||v||^2.
\end{aligned}$$

A condição (2.13) é chamada de **elipcidade uniforme do operador $\mathcal{L}$**. Agora, se $\mathcal{L}_t$, $t \in [0, T]$, for uma família de operadores diferenciais, diremos que a família é uniformemente elíptica se (2.13) acima se modifica para

$$\sum_{i,j=1}^{d} a_{i,j}(x)v_i v_j > L||v||, \qquad L > 0, \qquad \forall(t, x) \in [0, T] \times \mathrm{R}^d. \tag{2.14}$$

Uma última definição:

Definição 2.1.3. *Uma função $f : \mathrm{R}^d \to \mathrm{R}$ é dita Hölder contínua se*

$$|f(x) - f(y)| \leqslant C||x - y||^\alpha, \qquad \forall x, y \in \mathrm{R}^d,$$

onde C, α são costantes positivas.

Dessa forma, segundo (FRIEDMAN, 1964, Capítulo 1, Teoremas 12 e 15), temos a seguinte garantia de existência de solução para (2.9):

Teorema 2.1.4. *Considere o problema* (2.9)*. Suponha que a família de operaores* $\mathcal{L}_t$ *é uniformemente elíptica; que as funções* $a_{i,j}$*,* b_i *(da definição do operador* $\mathcal{L}$*) e* k *sejam limitadas; que* $a_{i,j}, b_j, k$ *e* g *sejam Hölder contínuas em* $[0,T] \times \mathrm{R}^d$*; e que* f *e* g *satifaçam as condições de crescimento polinomial enunciadas em (i) e (ii) do enunciado de 2.1.1. Então existe uma solução* v *para* (2.9) *satisfazendo a condição de crescimento* (2.10)

Observamos que a referência acima é fornecida, e as hipóteses levemente condensadas a fim de se enquadrar na nossa nomenclatura, por (KARATZAS; SHREVE, 1988).

Note que a condição de elipcidade uniforme e de limitação dos coeficientes a_{ij} pode parecer muito restritiva. Contudo, em muitas aplicações financeiras, a matriz σ (chamada de matriz de volatilidade) é constante e invertível, donde obtemos facilmente as hipóteses acima. Tendo demonstrado a representação de Feynman-Kac, o passo que anunciamos é a apresentação de um exemplo. Vamos usá-la, então, para resolver a equação do calor "backward".

Exemplo 2.1.5. *Considere o processo dado pelo movimento Browniano* n*-dimensional:*

$$dX_t = \sqrt{2\alpha} dW_t = \sqrt{2\alpha} I_n dW_t.$$

Então, $\sigma\sigma^T = 2\alpha I_n \Rightarrow (\sigma\sigma^T)_{i,j} = 2\alpha\delta_{i,j}$*. Assim,*

$$\mathcal{L}_t = \frac{1}{2}\sum_{i,j=1}^{n} 2\alpha\delta_{i,j}\partial_{i,j} = \alpha\sum_{i=1}^{n}\partial_i^2 = \alpha\Delta.$$

A equação do calor backward em n *dimensões,* $\alpha\Delta v - v_t = 0$*,* $v(T,x) = f(x)$*, assume a forma* (2.9)*, com* $k, g \equiv 0$*. Tomando* f *que satisfaz a condição de crescimento polinomial (i) de 2.1.1, concluimos que a equação tem solução e que tal solução é da forma*

$$v(t,x) = \mathrm{E}_{t,x}[f(W_T)], \qquad ,\forall (t,x) \in [0,T] \times \mathrm{R}^n.$$

Ou seja, a fim de saber o valor de $v(t,x)$*, iniciamos um movimento Browniano no instante* t *e na posição* x*, simulamos seu comportamento até o instante* T *e, então, encontramos o valor esperado de* f *nestes pontos.*

Antes de prosseguir com a representação de equações semilineares, vamos mostrar como podemos também representar problemas de contorno, como já vislumbrado em (2.8). Para estes problemas, contudo, não consideraremos mais EDPs parabólicas.

Teorema 2.1.6. *Seja D um domínio aberto e limitado de R^d, $k : \bar{D} \to [0, \infty)$, $g : \bar{D} \to \mathrm{R}$ e $f : \partial D \to \mathrm{R}$ funções contínuas, e $v : \bar{D} \to \mathrm{R}$ uma função contínua e de classe $C^2(D)$ satisfazendo o* ***problema de Dirichlet***

$$\begin{aligned} &\mathcal{L}v - kv = -g, \qquad x \in D; \\ &v = f, \qquad x \in \partial D. \end{aligned} \tag{2.15}$$

Se o tempo de parada $\tau_D = \inf\{t \geqslant 0; X_t \notin D\}$ for tal que $\mathrm{E}^x[\tau_D] < \infty$ para todo $x \in D$, então

$$v(x) = \mathrm{E}_x\Big[f(X_{\tau_D}) \exp\Big\{ - \int_0^{\tau_D} k(X_s)ds\Big\} + \int_0^{\tau_D} g(X_t) \exp\Big\{ - \int_0^t k(X_s)ds\Big\}dt\Big], \tag{2.16}$$

para todo $x \in D$.

Observe que, no enunciado desse teorema, estamos considerando o processo X_t que satisfaz

$$\begin{aligned} &dX_t = b(X_t)dt + \sigma(X_t)dW_t; \\ &X_0 = x, \end{aligned}$$

com o gerador infinitesimal $\mathcal{L}$ associado, e medida de probabilidade E_x associada ao movimento browniano iniciado em x. Antes da demonstração precisaremos de um lema técnico que pode ser encontrado com sua demonstração (insluive numa forma mais forte), em (MUNKRES, 1997, Lema 15.1, pág. 122).

Lema 2.1.7. *Seja D um aberto de R^n. Então existe uma sequência de compactos (K_n) contidos em D, cuja união é D, e tal que $K_n \subset Int(K_{n+1})$ para todo n.*

Demonstração. (do Teorema 2.1.6.) Defina $Y_s := \exp\Big\{ - \int_0^s k(X_s)ds\Big\}$. Seja (K_n) uma sequência de compactos como no lema 2.1.7 e considere os tempos de parada $\tau_n = \inf\{t \geqslant 0; X_t \notin K_n\}$. Considere o processo $Z_s = v(X_s)Y_s$. Então, fixado n, para qualquer $s \leqslant \tau_n$,

$$\begin{aligned} &dZ_s = Y_s(\mathcal{L}v(X_s)ds + Dv(X_s)\sigma(X_s)dW_s) - v(X_s)Y_s k(X_s)ds = \\ &Y_s((k(X_s)v(X_s) - g(X_s))ds + Dv(X_s)\sigma(X_s)dW_s) - v(X_s)Y_s k(X_s)ds \\ &= -Y_s g(X_s)ds + Y_s Dv(X_s)\sigma(X_s)dW_s, \end{aligned}$$

uma vez que, para tais valores de s, $X_s \in D$, e, portanto, usamos a EDP satisfeita por v. Assim, integrando dZ_s de 0 a s,

$$v(x) = v(X_s)Y_s + \int_0^s Y_u g(X_u) - \int_0^s Y_u Dv(X_u)\sigma(X_u)dW_u.$$

Substituindo no tempo de parada τ_n,

$$v(x) = v(X_{\tau_n})Y_{\tau_n} + \int_0^{\tau_n} Y_u g(X_u)du - \int_0^\infty 1_{\{u \leqslant \tau_n\}} Y_u Dv(X_u)\sigma(X_u)dW_u.$$

Vamos calcular a esperança E_x do último termo. Por simplicidade, considere o caso $d = 1$, de maneira que temos o integrando $1_{\{u \leqslant \tau_n\}} Y_u v'(X_u)\sigma(X_u)$. Como k é positivo, $0 \leqslant Y_s \leqslant 1$. Note que, por ser de classe C_2, existem constantes C_n tais que $|v'(x)| \leqslant C_n$, para todo $x \in K_n$. Assim, fixado $T > 0$, para qualquer $0 \leqslant u \leqslant T$,

$$|1_{\{u \leqslant \tau_n\}} Y_u v'(X_u)\sigma(X_u)|^2 \leqslant 1_{\{u \leqslant \tau_n\}} C_n^2 2K(1 + |X_u|^2) \leqslant 2C_n^2 K(1 + c_n^2),$$

onde c_n é tal que $K_n \subseteq B(c_n)$. Dessa maneira, para cada T, $1_{\{u \leqslant \tau_n\}} Y_u v'(X_u)\sigma(X_u) \in L^2([0,T] \times \Omega)$. Portanto, o processo

$$\left\{ \int_0^t 1_{\{u \leqslant \tau_n\}} Y_u v'(X_u)\sigma(X_u) dW_u \right\}_{t=0}^T$$

é um martingale pata todo T, e assim $\displaystyle\int_0^\infty 1_{\{u \leqslant \tau_n\}} Y_u v'(X_u)\sigma(X_u) dW_u$ tem esperança nula. Com isso,

$$v(x) = \mathrm{E}_x[v(X_{\tau_n})Y_{\tau_n} + \int_0^{\tau_n} Y_u g(X_u) du]. \tag{2.17}$$

Agora, notemos que $\tau_n \nearrow \tau_D$ (já que os compactos estão encaixados, sua união gera D, e o processo X_t é contínuo), de forma que

$$\begin{aligned} &v(X_{\tau_n})Y_{\tau_n} \to v(X_{\tau_D})Y_{\tau_D}, \text{ e} \\ &\int_0^{\tau_n} Y_u g(X_u) du \to \int_0^{\tau_D} Y_u g(X_u) du, \end{aligned}$$

por continuidade. Como v é contínua no compacto $\bar{D}$, temos $|v(X_{\tau_n})| \leqslant M$, para uma constante M e todo n. Assim, podemos usar o teorema da convergência dominada para concluir que

$$\lim_{n \to \infty} \mathrm{E}_x[v(X_{\tau_n})Y_{\tau_n}] = \mathrm{E}_x[v(X_{\tau_D})Y_{\tau_D})] = \mathrm{E}_x[f(X_{\tau_D})Y_{\tau_D})].$$

Ainda, seja N tal que $|g(x)| \leqslant N$ para todo $x \in \bar{D}$. Então

$$|\int_0^{\tau_n} Y_u g(X_u) du| \leqslant \int_0^{\tau_n} |g(X_u)| du \leqslant N\tau_n \leqslant N\tau_D,$$

e essa última variável aleatóra é integrável, por hipótese. Assim, usando o TCD,

$$\mathrm{E}_x[\int_0^{\tau_n} Y_u g(X_u) du] \to \mathrm{E}_x[\int_0^{\tau_D} Y_u g(X_u) du].$$

Tomando o limite $n \to \infty$ em (2.17) e usando a definição de Y_s, terminamos a demonstração. □

Mais uma vez, a representação de Feynman-Kac, em sua outra formulação, nos garante unicidade, sujeita à existência de solução. Novamente então gostaríamos de saber quando as condições do teorema 2.1.6 são satisfeitas. A primeira questão é sobre existência de uma solução para o problema. O próximo teorema resume um apanhado de condições encontradas em (KARATZAS; SHREVE, 1988) que garantem a existência de solução de (2.15). A demonstração pode ser encontrada, por exemplo, em (FRIEDMAN, 1964, pág. 87). Também temos as condições levemente modificadas para enquadrar nossa na nomenclatura. Antes do teorema precisamos de uma definição técnica.

Definição 2.1.8. *Dizemos aberto $D \subseteq \mathrm{R}^n$ satisfaz a condição de esfera exterior se, para todo $a \in \partial D$, existe uma esfera $B_x(y)$ em R^n tal que $B_x(y) \cap D = \varnothing$ e $B_x(y) \cap \partial D = \{a\}$.*

Teorema 2.1.9. *Suponha que o operador $\mathcal{L}$ seja uniformemente elíptico, as funções $a_{i,j}, b_j$ (que definem $\mathcal{L}$), k e g são Hölder-contínuas e D satisfaz a condição de esfera exterior. Então existe uma função $v \in C^2(D) \cap C(\bar{D})$ que resolve (2.15).*

A segunda questão sobre as hipóteses do Teorema 2.1.6 é sobre quando o tempo de parada τ_D possui esperança finita. O próximo teorema, também vindo de (KARATZAS; SHREVE, 1988), versa sobre isso.

Teorema 2.1.10. *Suponha que para algum $1 \leqslant l \leqslant d$, nos tenhamos*

$$\inf_{x \in \bar{D}} a_{ll}(x) > 0. \tag{2.18}$$

Então, $\mathrm{E}_x(\tau_D) < \infty$, $\forall x \in D$.

Demonstração. Vamos definir as constantes $b := \sup_{x \in \bar{D}} |b(x)|$; $a := \inf_{x \in \bar{D}} a_{ll}(x)$; $q := \inf_{x \in \bar{D}} x_l$, possível por continuidade e compacidade de $\bar{D}$. Agora, seja $v > 2b/a$. Tome a função $h : D \to \mathrm{R}$, $h(x) = -\mu e^{vx_l}$. Note que a função h é C^∞ em D e

$$\partial_i h(x) = \begin{cases} 0, \text{se } i \neq l, \\ -\mu v e^{vx_l}, \text{se } i = l \end{cases} \quad \Rightarrow \partial_{ij} h(x) = \begin{cases} 0, \text{se } (i,j) \neq (l,l), \\ -\mu v^2 e^{vx_l}, \text{se } (i,j) = (l,l). \end{cases}$$

Dessa forma,

$$\begin{aligned} &-(\mathcal{L}h)(x) = -(\frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(x)\partial_{ij}h(x) + \sum_{i=1}^{d} b_i(x)\partial_i h(x)) = \\ &-(-\frac{1}{2}a_{ll}(x)\mu v^2 e^{vx_l} - b_l(x)\mu v e^{vx_l}) = \mu v e^{vx_l}(\frac{1}{2}v a_{ll}(x) + b_l(x)) \geqslant \\ &\frac{1}{2}\mu v e^{vq} a(v + \frac{2b_l(x)}{a}) \geqslant \frac{1}{2}\mu v e^{vq} a(v - \frac{2b}{a}); \qquad x \in D. \end{aligned} \tag{2.19}$$

Sendo $a > 0$ por hipótese e pela escolha de v, podemos escolher μ grande a ponto de

$$\mathcal{L}h(x) \leqslant -\frac{1}{2}\mu v e^{vq} a(v - \frac{2b}{a}) \leqslant -1, \qquad x \in D.$$

Agora, pela fórmula de Itô,

$$dh(X_t) = \mathcal{L}h(X_t)dt + Dh(X_t)\sigma(X_t)dW_t,$$

de onde,

$$h(X_{t\wedge\tau_D}) = h(x) + \int_0^{t\wedge\tau_D} \mathcal{L}h(X_s)ds + \int_0^t 1_{\{u\leqslant\tau_D\}} Dh(X_s)\sigma(X_s)dW_s.$$

Notando que h, Dh e σ são limitados em $\bar{D}$, concluímos que o último processo é um martingale. Tomando esperança, lembrando que $h(x)$ é deterministico e notando que $|h|$ é limitado por, digamos, M em D, temos

$$\mathrm{E}[h(X_{t\wedge\tau_D})] = h(x) + \mathrm{E}[\int_0^{t\wedge\tau_D} \mathcal{L}h(X_s)ds] \leqslant h(x) + \mathrm{E}[\int_0^{t\wedge\tau_D} -1ds] \Rightarrow$$
$$\mathrm{E}[\tau\wedge t] \leqslant h(x) - \mathrm{E}[h(X_{t\wedge\tau_D})] \Rightarrow \mathrm{E}[\tau_D\wedge t] \leqslant |h(x)| + \mathrm{E}[|h(X_{t\wedge\tau_D})|] = 2M.$$

Assim, pelo lema de Fatou, lembrando que a sequência $t\wedge\tau_D$ é não negativa,

$$\mathrm{E}[\tau_D] = \mathrm{E}[\liminf_{t\to\infty} \tau_D\wedge t] \leqslant \liminf_{t\to\infty} \mathrm{E}[t\wedge\tau_D] \leqslant 2M.$$

O resultado segue. □

O resultado acima, que fornece condição suficiente sob a qual uma das hipóteses de 2.1.6 é satisfeita, tem o corolário imediato:

Corolário 2.1.11. *Seja W_t um movimento Browniano em d-dimensões, e seja D um aberto limitado de R^d. Então, W_t escapa de D em tempo finito.*

Demonstração. De fato, o processo é representado diferencialemnte por $dX_t = I_n dW_t$, de forma que a matriz $A = \sigma\sigma^T = I$ satisfaz trivialmente as hipóteses de 2.1.10 □

Vamos mostrar alguns exemplos/aplicações do Teorema 2.1.6.

Exemplo 2.1.12. *Seja D um domínio aberto, limitado e cm a propriedade de esfera exterior. Suponha que $\mathcal{L}$ seja um operador elíptico em D e que $f : \partial D \to \mathrm{R}$ seja contínua. Então, o problema de Dirichlet*

$$\mathcal{L}v = 0, \qquad \forall x \in D;$$
$$v = f, \qquad \forall x \in D.$$

tem solução pelo teorema 2.1.9. Mais ainda, por 2.1.6, v tem a forma

$$v(x) = \mathrm{E}_x[f(X_{\tau_D})].$$

Em particular, notamos que o valor de v independe o caminho traçado pelo processo X ou do tempo que ele demora para atingir a fronteira ∂D. Depende apenas onde de ele a atinge.

Exemplo 2.1.13. *Do exemplo anterior, 2.1.12, considerando o movimento Browniano W_t em n-dimensões e lembrando que, neste caso, o gerador infinitesimal $\mathcal{L}$ é o laplaciano ∇, temos a seguinte representação paraa* ***equação de Laplace****, $\nabla v = 0$ para $x \in D$ e $v(x) = f(x)$ para $x \in \partial D$:*

$$v(x) = \mathrm{E}_x[f(W_{\tau_D})],$$

e obtemos uma relação entre funções harmônicas e o movimento Browniano.

Lembremos agora que há um corolário do princípio do máximo para funções harmônicas v que diz que se v for harmônica em D e constante no bordo de ∂D, então v é constante em D. Isso é facilmente observado em 2.1.13. Notemos agora que o exemplo 2.1.12 generaliza esse resultado para operadores elípticos quaisquer, desde que possam ser escritos como geradores infinitesimais de processos de Itô X_t.

Exemplo 2.1.14. *Tomando $g = -1$, $k = 0$ e $f = 0$ em 2.15, obtemos*

$$v(x) = \mathrm{E}_x[\tau_D].$$

Isso nos permite encontrar o valor esperado do tempo τ_D que o processo X_t, iniciado em x demora para atindir a fronteira ∂D, através da resolução de uma EPD.

Exemplo 2.1.15. *Condiere agora o processo equivalente a um movimento Browniano padrão em dimensão 1. O operador $\mathcal{L}$ associado é, então, $\frac{d^2}{dx^2}$. Seja $k = 0$, $g = 0$ e $f(x) = x$. Considere o aberto na reta $D = (-a, b)$, $a, b > 0$. O problema de Diriclhet associado é*

$$\frac{d^2 v}{dx^2} = 0, \qquad x \in (a, b);$$
$$v(-a) = -a; v(b) = b.$$

A solução desse problema é facilmente dada por $v(x) = x$ (pois tomamos $a + b > 0$). Pela representação 2.16, e notando que $W_{\tau_D} = -a$ ou b, temos

$$0 = \mathrm{E}[W_{\tau_D}] = -a\mathrm{P}(W_{\tau_D} = -a) + b\mathrm{P}(W_{\tau_D} = b).$$

Ainda,

$$\mathrm{P}(W_{\tau_D} = -a) + \mathrm{P}(W_{\tau_D} = b) = 1.$$

Dessa forma, $P(W_{\tau_D} = -a) = a/(a + b)$ e $P(W_{\tau_D} = b) = b/(a + b)$. Na prática, isso nos diz que se uma partícula (ou uma ação, por exemplo) estiver se movendo de acordo com um movimento Browniano, decidirmos pará-la quando atingir os patamares $-a$ ou b e estivermos interessados em entender em qual dos valores ela tem mais probabilidade se estacionar, o cálculo é feito como acima, e o resultado é conhecido. Isso nos mostra uma propriedade do movimento Browniano que reflete o mesmo comportamento observado em passeios aleatórios. Uma outra abordagem do mesmo problema é feita em (STEELE, 2001, Capítulo 1).

2.2 Feynman-Kac para equações semilineares

Seguindo a sequência prometida na seção anterior, em que intrduzimos representações estocásticas para EDPs lineares do tipo

$$\mathcal{L}_t v(t,x) + \frac{dv(t,x)}{dt} - kv(t,x) = -g(t,x),$$

sujeitas à condição terminal

$$v(x,T) = f(x),$$

e mostramos alguns exemplos e aplicações, chegamos à parte de generalizar o resultado. Vamos considerar então EDPs **semilineares** que são do tipo

$$\mathcal{L}_t v(t,x) + \partial_t v(t,x) + f(t,x,v(t,x),D_x v(t,x)) = 0, \tag{2.20}$$

onde mantemos a notação da seção anterior, $\mathcal{L}_t$ gerador infinitesimal do processo (2.4). Como exemplos desse tipo de EDP, temos

$$\begin{aligned}
-\partial_t S = H(q,\partial_q S,t) &\qquad \text{Equação de Hamilton-Jacobi;}\\
\partial_t u = D\partial_x^2 u + R(u) &\qquad \text{Equações de reação-difusão;}\\
\partial_t u - d\partial_x^2 u = u\partial_x u &\qquad \text{Equação de Burger.}
\end{aligned}$$

Como na outra seção, primeiramente vamos tentar encontrar, de forma ingênua, uma representação para uma solução de da EDP acima. Seguiremos aqui os passos da segunda parte da referência (PERKOWSKI, 2011). Suponha então que v seja uma solução suficientemente suave da EDP 2.20, e não nos preocupamos inicialmente com condições terminais ou iniciais. Considere uma versão do processo X_t, que lembramos ser dado pela SDE,

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dW_t,$$

iniciado no instante t na posição x. Denotemos o processo por $X^{t,x}$. Denote por $Y^{t,x}$ o processo $Y_s^{t,x} := v(x, X_s^{t,x})$. Com isso, somos intuídos a tentar entender a dinâmica do processo $Y^{t,x}$ uma vez que seu valor, no tempo t, é $v(t, X_t^{t,x}) = v(t,x)$. Abreviaremos doravante $X^{t,x}$ por X e $Y^{t,x}$ por Y. Note que, pela fórmula de Itô e pela hipótese em v,

$$\begin{aligned}
dY_s &= (\partial_s v(s,X_s) + \mathcal{L}_s v(s,X_s))ds + D_x v(s,X_s)\sigma(s,X_s)dW_s =\\
&- f(s,X_s,v(s,X_s),D_x v(s,X_s))ds + D_x v(s,X_s)\sigma(s,X_s)dW_s =\\
&- f(s,X_s,Y_s,D_x v(s,X_s))ds + D_x v(s,X_s)\sigma(s,X_s)dW_s.
\end{aligned}$$

Com isso, a SDE que descreve a dinâmica de Y_s possui duas apresentações possivelmente distintas de um precesso desconhecido, $D_x v(s,X_s)$ e $D_x v(s,X_s)\sigma(s,X_s)$. Consideramos então EDPs semilineares um pouco menos gerais, da forma

$$\mathcal{L}_t v(t,x) + \partial_t v(t,x) + f(t,x,v(t,x),D_x v(t,x)\sigma(t,x)) = 0. \tag{2.21}$$

Note que quando $\sigma(t,x) \in R^{d\times d}$ é invertível para todo par (t,x), as expressões (2.20) e (2.21) descrevem a mesma classe de problemas. Basta tomar em (2.21) $\tilde{f}$ em vez de f, onde

$$\tilde{f}(\cdot,z) = f(\cdot, z\sigma^{-1}(t,x)).$$

Considerando apenas equações da forma (2.21) e definindo o processo $Z_s := D_x v(s,X_s)\sigma(s,X_s)$, temos que a dinâmica de Y satisfaz

$$dY_s = -f(s,X_s,Y_s,Z_s)ds + Z_s dW_s. \tag{2.22}$$

A equação acima possui dois processos incógnitos, e uma solução para ela seria um par de processos (Y_s, Z_s). Lembremos que estamos estamos interessados em encontrar, com unicidade, a dinâmica do processo Y_s. Com unicidade pois, se a equação (2.22) possuísse mais de uma solução em Y_s, não teríamos como relacioná-las com $v(t, X_t^{t,x})$. Voltando à (2.22), note que ele é equivalente a

$$Y_s = Y_0 - \int_0^t f(s,X_s,Y_s,Z_s)ds + \int_0^t Z_s dW_s.$$

Tome então o caso particular $f \equiv 0$. Temos

$$Y_s = Y_0 + \int_0^t Z_s dW_s,$$

de forma que, para cada processo Z_s integrável segundo Itô, temos uma solução distinta para Y_s. Com isso, não podemos esperar resolver satisfatoriamente o problema (2.22) se nos for fornecida uma **condição inicial** Y_0. Suponha então que, junto da equação (2.22), nos é fornecida a **condição terminal** $Y_T = \xi$, onde ξ é uma variável aleatória $\mathcal{F}_T$ mensurável. Então, integrando (2.22) de t a T, ainda no caso particular $f \equiv 0$,

$$Y_T = Y_t + \int_t^T Z_s dW_s \Rightarrow Y_t = \xi - \int_t^T Z_s dW_s. \tag{2.23}$$

Tomando a esperança condicional em ambos os lados,

$$Y_t = E[Y_t|\mathcal{F}_t] = \mathrm{E}[\xi - \int_t^T Z_s dW_s|\mathcal{F}_t] = \mathrm{E}[\xi|\mathcal{F}_t] - \mathrm{E}[\int_t^T Z_s dW_s] = \mathrm{E}[\xi|\mathcal{F}_t],$$

onde a última igualdade segue do fato da integral de Itô ser um martingale para Z_t adequado e da independência dos incrementos. Isso nos diz que Y_t, dadas condições adequadas de integrabilidade de ξ, é um martingale. Supondo que estejamos considerando, no espaço de probabilidade, a filtração browniana, podemos usar o teorema de representaçao de martingales para escrever, **de forma única**,

$$Y_t = Y_0 + \int_0^t Z_s dW_s,$$

para algum procsso Z_s. Assim, por (2.23),

$$Y_t = \xi - \int_t^T Z_s dW_s,$$

isto é, partindo apenas da condição terminal e da expressão da equação, encontramos, **unicamente**, um par de processos (Y, Z) que satisfazem a equação estocástica e a condição terminal simultaneamente. Claro é, contudo, que a noção de unicidade precisa ser mais bem estabelecida, por exemplo, precisa-se saber em que espaços procuramos soluçõs. Entretanto, já podemos vislumbrar de que sorte são os problemas que faz sentido considerarmos. Estudaremos então equações estocásticas do tipo

$$Y_t = f(t, Y_t, Z_t)dt + Z_t dW_t, \qquad t \in [0, T]; \tag{2.24}$$

$$Y_T = \xi \in \mathcal{F}_T. \tag{2.25}$$

Essa derivação heurística de um problema estocástico que surge através da EDP semilinear (2.20) gera algumas questões: como exatamente a solução de (2.20) se relaciona com a dinâmica do processo Y_t que é solução de (2.24), (2.25); o que exatamente é uma solução de (2.24); quando há existência e/ou unicidade de soluções de (2.24), (2.25); em que espaços procuramos essas soluções. Essa última questão é importante, por exemplo, para a propriedade de martingale da integral de Z_s que utilizamos acima. Responderemos essas questões nas duas próximas seções.

2.3 BSDEs - Definição, Existência e Unicidade de soluções.

Nesta seção, vamos seguir o desenvolvimento da teoria de BSDEs dada por (HU, 2013, parte 1). Uma equação diferencial estocástica "backward", abreviada por BSDE (do inglês Backward Stochastic Differential Equation), é uma equação e uma condição terminal do tipo

$$Y_t = -f(t, Y_t, Z_t)dt + Z_t dW_t, \qquad t \in [0, T]; \tag{2.26}$$

$$Y_T = \xi \in \mathcal{F}_T, \tag{2.27}$$

precisamente como em (2.24), (2.25), mas com um sinal negativo na frente de f a fim de simplificar a notação. É justamente por ser apresentada em conjunto com uma condição terminal (já que não podemos esperar unicidade de solução para equações do tipo (2.26) com uma condição inicial) que chamamos a equação de backward. Um par de processos (Y, Z) é dito solução de (2.26), (2.27) se

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \qquad t \in [0, T]. \tag{2.28}$$

Antes de estudar resultados sobre existência e unicidade, vamos fixar os espaços de processos onde buscaremos as soluções. Primeiro, uma questão de notação. Usaremos nesta seção $X(Y)$ (respectivamente $X[Y]$) para enfatizar um conjunto X de funções com domínio (respectivamente contradomínio) em Y.

Definição 2.3.1. *1. Denotamos por $M^2[\mathrm{R}^k]$ o espaço dos processos $Y : \Omega \times [0,T] \to \mathrm{R}^k$ progressivamente mensuráveis tais que*

$$\mathrm{E}[\sup_{t\in[0,T]} |Y_t|^2] < \infty.$$

2. Denotamos por $M_c^2[\mathrm{R}^k]$ o subconjunto de $M^2[\mathrm{R}^k]$ cujos processos são contínuos.

3. Denotamos por $L^2[\mathrm{R}^{k\times n}]$ o conjunto dos processos $Z : \Omega \times [0,T] \to \mathrm{R}^{k\times n}$ tais que

$$\mathrm{E}[\int_0^T ||Z(t)||^2 dt] < \infty,$$

onde consideramos a norma $||z||^2 = tr\{zz^T\}$ em $\mathrm{R}^{k\times n}$.

Obsevações:

- Omitiremos, a menos quando imprescindível, os argumentos R^k e $\mathrm{R}^{k\times n}$ de $M^2[\mathrm{R}^k]$ e $L^2[\mathrm{R}^{k\times n}]$.
- $\mathrm{E}[\sup_{t\in[0,T]} |Y_t|^2]$ e $\mathrm{E}[\int_0^T ||Z(t)||^2 dt]$ definem normas (ao quadrado) em M^2 e L^2, respectivamente.
- M^2, M_c^2 e L^2 são espaços de Banach com a norma acima.
- Note que $M_c^2[\mathrm{R}^k] \cap \{\text{martingales}\} = L^2[\mathrm{R}^k] \cap \{\text{martingales quadrado integráveis}\}$, pela desigualdade de Doob A.3.3.

Antes de estudar o BSDE (2.26), (2.27), precismos de condições mínimas sobre f e ξ para definir (2.26), além de nomenclatura.

1. Reforçamos que estamos trabalhando com um espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$ fixado, onde temos definido um movimento browniano n-dimensional padrão, W_t, e consideramos $\mathcal{F}_t$ a filtração browniana neste espaço.

2. Consideramos um mapa

$$f : \Omega \times [0,T] \times \mathrm{R}^k \times \mathrm{R}^{k\times n} \to \mathrm{R}^k$$

$\mathcal{F} \otimes \mathcal{B}([0,T]) \otimes \mathcal{B}(\mathrm{R}^k) \otimes \mathcal{B}(\mathrm{R}^{k\times n})$-mensurável, tal que, para cada $y \in \mathrm{R}^k$ e $z \in \mathrm{R}^{k\times n}$, $f(\cdot,\cdot,y,z)$ é progressivamente mensurável. O caso mais comum é quando $f(\omega,\cdot,\cdot,\cdot)$ é contínua para todo ω.

3. $\xi : \Omega \to \mathrm{R}^k$ é $\mathcal{F}_T$-mensurável.

4. Chamamos f de **gerador** do BSDE e ξ de **condição terminal**.

5. Costumamos omitir a dependência em ω de f, escrevendo, geralmente, $f(t, y, z)$ em vez de $f(\omega, t, y, z)$ e habituamos entender as afirmações de maneira qtp.

Definição 2.3.2. *Uma solução pro BDSE* (2.26), (2.27) *é um par de processos* (Y, Z), $Y : \Omega \times [0, T] \to \mathrm{R}^k$, $Z : \Omega \times [0, T] \to \mathrm{R}^{k \times n}$ *progressivamente mensuráveis tais que*

1. $\displaystyle\int_0^T |f(t, Y_t, Z_t)|dt < \infty$, *qtp*

2. $\displaystyle\int_0^T |Z_t|^2 dt < \infty$, *qtp.*

3. $$Y_t = \xi + \int_t^T f(r, Y_r, Z_r)dr - \int_t^T Z_r dW_r, \qquad \forall t \in [0, T].$$

Note que:

- Pela mensurabilidade de f, Y e Z, além da condição 1 da definição acima, a primeira integral em 3 é bem definida.

- Pela mensurabilidade de Z_t e a condição 2 da definição acima, a integral estocástica em 3 é bem definida.

- Por ser expresso como integrais de processos, Y_t é contínuo.

- O requerimento de Y_t ser adaptado implica que, necessariamente, Y_0 é determinístico e portanto uma constante (já que $\mathcal{F}_0 = \{\emptyset, \Omega\}$). Isso tem uma consequência direta na precificação de derivativos, que veremos na seção sobre o modelo de Black-Scholes.

A princípio, não pedimos nenhuma condição de integrabilidade sobre o processo Y_t que faz parte da solução BSDE. E isso porque o próximo resultado **nos fornece** boas condições de integrabilidade em Y_t, quando este for solução do BSDE gerado por um processo f que satisfaz condições razoáveis de crescimento. Antes, um lema técnico bem difundido no estudo de equações diferenciais.

Lema 2.3.3 (Grownall)**.** *Sejam* $\phi \to \mathrm{R}$ *uma função limitada, não negativa e (borel) mensurável,* $B > 0$ *e* $C : [0, T] \to \mathrm{R}$ *uma função integrável e não negativa tais que*

$$\phi(t) \leqslant B + \int_0^t C(r)\phi(r)dr, \qquad \forall t \in [0, T].$$

Então

$$\phi(t) \leqslant B \exp\Big(\int_0^t C(r)dr\Big), \qquad \forall t \in [0, T].$$

Demonstração. (TESCHL, 2012, Lema 2.7) □

Proposição 2.3.4. *Suponha que exista um processo $G \in L^2(\Omega \times [0,T])$ e uma constante c tais que*

$$|f(t,y,z)| \leqslant |G_t| + c(|y| + |z|), \qquad \forall (t,y,z) \in [0,T] \times \mathrm{R}^k \times \mathrm{R}^{k \times n}.$$

Então, se (Y,Z) for uma solução do BSDE (2.26), (2.27) tal que $Z \in L^2$, teremos $Y \in M_c^2$.

Demonstração. Integrando (2.26) de 0 a t, temos

$$Y_t = Y_0 - \int_0^t f(r, Y_r, Z_r)dr + \int_0^t Z_r dW_r,$$

de onde,

$$|Y_t| \leqslant (|Y_0| + \int_0^T (|G_r| + c|Z_r|)dr + \sup_{t \in [0,T]} |\int_0^t Z_r dW_r|) + \int_0^t c|Y_r|dr.$$

Agora, notemos que

1. $|Y_0|$ é uma constante, portanto limitada;
2. $G_t \in L^2(\Omega \times [0,T]) \Rightarrow G_t \in L^1(\Omega \times [0,T])$, já que temos um espaço de medida finito. Assim, $\mathrm{E}[\int_0^T |G_t|dt] < \infty \Rightarrow \int_0^T |G_t|dt < \infty$ qtp.
3. $\int_0^T |Z_t|dt < \infty$ qtp, pelo mesmo raciocínio anterior.
4. Como $Z_t \in L^2$, a integral estocástica $\int_0^t Z_r dW_r$ é um martingale, de forma que, aplicando a desigualdade de Doob A.3.3,
$$\mathrm{E}\Big[\sup_{t \in [0,T]} |\int_0^t Z_r dW_r|\Big] \leqslant k\mathrm{E}|\int_0^T Z_t dW_t|.$$
Pela isometria de Itô e por A.1.4,
$$\mathrm{E}|\int_0^T Z_t dW_t| \leqslant \mathrm{E}|\int_0^T Z_t dW_t|^2 = \mathrm{E}[\int_0^T Z_t^2 dt] < \infty.$$
5. Y_t é limitado qtp por continuidade.

Assim, podemos definir a variável aleatória

$$B(\omega) = |Y_0| + \int_0^T (|G_r(\omega)| + c|Z_r(\omega)|)dr + \sup_{t \in [0,T]} |\int_0^t Z_r(\omega) dW_r(\omega)|,$$

e pela desigualdade de Gronwall,

$$|Y_t(\omega)| \leqslant B(\omega) \exp\Big(\int_0^t c\Big) \leqslant B(\omega) \exp(cT),$$

de onde,

$$\sup_{t\in[0,T]} |Y_t(\omega)|^2 \leqslant CB(\omega)^2,$$

com C constante. Note que $\mathrm{E}[B^2] < \infty$, já que

1. $|Y_0|$ é constante;
2. $\mathrm{E}[(\int_0^T |G_r|dr)^2] \leqslant T\mathrm{E}[\int_0^T |G_t|^2 dt] < \infty$;
3. $\mathrm{E}[(\int_0^T |Z_r|dr)^2] \leqslant T\mathrm{E}[\int_0^T |Z_t|^2 dt] < \infty$;
4. Pela desigualdade de Doob e Isometria de Itô,
$$\mathrm{E}\Big[\sup_{t\in[0,T]} |\int_0^t Z_r dW_r|^2\Big] \leqslant C_2\mathrm{E}\Big[\int_0^T Z_t^2 dt\Big] < \infty.$$

Dessa forma, concluimos que $\mathrm{E}[\sup_{t\in[0,T]} |Y_t|^2] < \infty$, e $Y \in M_c^2$. □

Agora, um lema técnico.

Lema 2.3.5. *Sejam $Y_t \in M^2[\mathrm{R}^k]$ e $Z_t \in L^2[\mathrm{R}^{k\times n}]$. Então o processo*

$$\int_0^t Y_s \cdot Z_s dW_s, \qquad t \in [0,T]$$

é um ***martingale*** *e uniformemente integravél, isto é,*

$$\mathrm{E}[|\sup_{t\in[0,T]} \int_0^t Y_s \cdot Z_s dW_s|] < \infty.$$

Demonstração. Denote por $G_s = (G_s^1, \ldots, G_s^n)$ o processo em R^n $Y_s Z_s$. Então,

$$\mathrm{E}[\sup_{t\in[0,T]} |\int_0^t G_s dW_s|] = \mathrm{E}[\sup_{t\in[0,T]} |\sum_{i=1}^n \int_0^t G_s^i dW_s^i|].$$

Denotemos por H_s^i o processo $\int_0^t G_s^i dW_s^i$. Note que H_s^i é um martingale local, já que temos um integrando contínuo na integral de Itô (e é bem por isso que ela é definida). Então, pela de desigualdade de Burkholder-Davis-Gundy,

$$\mathrm{E}[\sup_{t\in[0,T]} |\sum_{i=1}^n H_t^i|] \leqslant C\mathrm{E}[\langle \sum_{i=1}^n H^i \rangle_T^{1/2}] = C\mathrm{E}[(\sum_{i,j=1}^n \langle H^i, H^j\rangle_T)^{1/2}].$$

Agora,

$$\langle H^i, H^j\rangle_T = \langle I_{W^i}(G^i), I_{W^j}(G^j)\rangle_T = \int_0^T G_s^i G_s^j d\langle W^i, W^j\rangle_s = 0, \; i \neq j$$

$$\langle H^i\rangle_T = \int_0^T (G_t^i)^2 dt.$$

Com isso,

$$\mathrm{E}[\sup_{t\in[0,T]}|\sum_{i=1}^{n}H_t^i|] \leqslant C\mathrm{E}[(\sum_{i=1}^{n}\int_0^T(G_t^i)^2dt)^{1/2}] = C\mathrm{E}[(\int_0^T\sum_{i=1}^{n}(G_t^i)^2dt)^{1/2}]$$

Chamando de Z_t^i a i-ésima coluna da matriz Z_t e denotando por $\cdot$ o produto interno do R^k, temos

$$G_t^i = Y_t \cdot Z_t^i \Rightarrow (G_t^i)^2 \leqslant |Y_t|^2|Z_t^i|^2,$$

normas de R^k. Logo,

$$\sum_{i=1}^{n}(G_t^i)^2 \leqslant |Y_t|^2\sum_{i=1}^{n}(Z_t^i)^2 = |Y_t|^2|Z_t|^2,$$

onde estamos considerando a norma do traço para Z_t. Assim

$$\mathrm{E}[\sup_{t\in[0,T]}|\int_0^t G_s dW_s|] \leqslant C\mathrm{E}[(\int_0^T|Y_t|^2|Z_t|^2)^{1/2}dt] \leqslant$$

$$C\mathrm{E}[\sup_{t\in[0,T]}|Y_t|(\int_0^T|Z_t|^2dt)^{1/2}] \leqslant C'\mathrm{E}[\sup_{t\in[0,T]}|Y_t|^2] + \mathrm{E}[\int_0^t|Z_t|^2dt] < \infty,$$

onde usamos o fato de que $2ab \leqslant a^2 + b^2$ e as hipóteses sobre Y e Z. Ora, sendo o processo $\int_0^t Y_s \cdot Z_s dW_s$ um martingale local uniformemente integrável, concluimos que ele é um martingale (por A.3.5), e o resultado segue.

□

Agora, seguiremos em direção de mostrar um teorema de existência e unicidade para soluções do BDSE (2.26), (2.27). Buscaremos soluções no espaço $M_c^2(\mathrm{R}^k)\times L^2(\mathrm{R}^{k\times n})$. Precisaremos de algumas condições sobre o gerador f; como é de se esperar dentro do contexto de equações diferenciais, analisaremos o caso em que f é Lipschitz. Esse resultado aparece primeiramente no artigo de Pardoux e Peng (PARDOUX; PENG, 1990). Antes dele, vale notar que já se existia um resultado conhecido para o caso em que f é linear, encontrado por inspeção. Comecemos enunciando as hipóteses completas sobre o gerador f e a condição terminal ξ.

Definição 2.3.6. *Diremos que o BSDE (2.26), (2.27) satisfaz as **hipóteses usuais**, se*

1. *$f : \Omega\times[0,T]\times\mathrm{R}^k\times\mathrm{R}^{k\times n}\to\mathrm{R}^k$ é mensurável e, para cada $(y,z)\in\mathrm{R}^k\times\mathrm{R}^{k\times n}$, $f(\cdot,\cdot,y,z)$ é progressivamete mensurável.*
2. *ξ é uma variável aleátoria $\mathcal{F}_T$ mensurável.*
3. *Existe uma constante $L > 0$ tal que para todo t, y, y', z, z' (em seus respectivos espaços óbvios), temos*

$$|f(t,y,z) - f(t,y',z')| \leqslant L(|y-y'| + |z-z'|),\ \mathrm{P}\ \textit{quase sempre},$$

onde consideramos as normas apropriadas já discutias nos espaços R^{algo}.

4. *(Integrabilidade)*

$$\mathrm{E}[|\xi|^2 + \int_0^T |f(t,0,0)|^2 dt] < \infty.$$

A ideia da demonstração é resolver o BSDE para geradores que independem de (y, z) e depois adaptar esse resultado para geradores com essa dependência. E isso porque, na primeira etapa, usamos um resultado sobre representação de martingales, tendo um olhar mais "probabilístico" do problema. Talvez seja o grande núcleo da questão. Já na segunda, usamos o teorema do ponto fixo de Banach, e temos uma abordagem mais analítica. Podemos entender essa etapa como uma lapidação do resultado anterior.

Teorema 2.3.7. *Seja* $\xi \in L^2(\Omega, \mathcal{F}_T)$ *e* F_t *um processo em* $L^2(\Omega \times [0,T])$*. Então, o BSDE*

$$\begin{aligned} dY_t &= -F_t dt + Z_t dW_t, \qquad t \in [0,T] \\ Y_T &= \xi, \end{aligned}$$

tem uma única solução no espaço $M_c^2[\mathrm{R}^k] \times L^2[\mathrm{R}^{k\times n}]$*, isto é, existe um único par de processos* $(Y_t, Z_t) \in M_c^2[\mathrm{R}^k] \times L^2[\mathrm{R}^{k\times n}]$ *tal que*

$$Y_t = \xi + \int_t^T F_s ds - \int_t^T Z_s dW_s, \qquad \forall t \in [0,T]. \tag{2.29}$$

Demonstração. Como procuramos uma solução para o BSDE acima? Vamos começar suponho que existe uma no espaço desejado e entender sua forma. Tomando esperança condicional em relação a $\mathcal{F}_t$ em (2.29),

$$Y_t = \mathrm{E}[\xi + \int_t^T F_s ds | \mathcal{F}_t] = \mathrm{E}[\xi + \int_0^T F_s ds | \mathcal{F}_t] - \int_0^t F_s ds, \tag{2.30}$$

onde usamos que $\int_t^T Z_s dW_s$ é independente de $\mathcal{F}_t$ e que sua esperança é nula, por ser um martingale. Assim, pela forma (2.30) de Y acima, **definimos** o processo Y_t como

$$Y_t = \mathrm{E}[\xi + \int_0^T F_s ds | \mathcal{F}_t] - \int_0^t F_s ds = M_t - \int_0^t F_s ds. \tag{2.31}$$

Perceba que, por hipótese sobre ξ e $\mathcal{F}$, os processos dentro da esperança condicional acima são integráveis (a norma 1 é limitada pela norma 2), de maneira que o processo gerado pelas esperanças condicionais $\mathrm{E}[\cdot|\mathcal{F}_t]$ forma uma martingale M_t. Agora, lembrando que $\mathcal{F}_t$ é a filtração browniana, podemos usar o teorema de represeNtação de martingales para encontrar um único processo Z_t em $L^2[\mathrm{R}^{k\times n}]$ tal que

$$M_t = M_0 + \int_0^t Z_s dW_s. \tag{2.32}$$

Agora, note que (Y, Z) assim dado, Y por (2.31) e Z por (2.32), formam uma solução do BSDE (2.29) como veremos. Primeiramente,

$$Y_T = \mathrm{E}[\xi + \int_0^T F_s ds | \mathcal{F}_T] - \int_0^T F_s ds = \xi,$$

pelas hipóteses de mensurabilidade. E, com isso,

$$Y_t - \xi = Y_t - Y_T = (M_0 + \int_0^t Z_s dW_s - \int_0^t F_s ds) - (M_0 + \int_0^T Z_s dW_s - \int_0^T F_s ds)$$
$$= \int_t^T F(s) ds - \int_t^T Z_s dW_s,$$

o que mostra que o par (Y, Z) satisfaz (2.29). Agora, (Y, Z) é uma solução pro BSDE tal que $Z \in L^2$, e concluimos, pela proposição 1, uma vez que F satisfaz suas hipóteses, que $Y_t \in M_c^2$. Temos, portanto, uma solução para o BDSE no espaço desejado. A unicidade vem da observação de que se (Y, Z) for outra solução neste espaço, Y necessariamente tem que ter a forma (2.31), e Z necessariamente será o único processo satisfazendo (2.32). O resultado segue.

□

Teorema 2.3.8 (Pardoux-Peng)**.** *Sob as hipóteses usuais, o BSDE* (2.26)*,* (2.27) *tem uma solução única em* $M_c[\mathrm{R}^k] \times L_2[\mathrm{R}^{k \times n}]$.

Demonstração. Seja $\Phi : M_c[\mathrm{R}^k] \times L_2[\mathrm{R}^{k\times n}] \to M_c[\mathrm{R}^k] \times L_2[\mathrm{R}^{k\times n}]$ tal que $\Phi(U, V) = (Y, Z)$ é a única solução do BSDE

$$dY_t = -f(t, U_t, V_t)dt + V_t dW_t \tag{2.33}$$
$$Y_T = \xi. \tag{2.34}$$

Note que o processo $F(\omega, t) := f(\omega, t, U(\omega, t), V(\omega, t))$ pertence a $L^2(\Omega \times [0, T])$, já que

$$|F_t| = |f(t, U_t, V_t) - f(t, 0, 0) + f(t, 0, 0)| \leqslant L(|U_t| + |V_t|) + |f(t, 0, 0)|,$$

e esses últimos processos estão em L^2, por hipótese. Logo, pelo teorema anterior, há uma única solução de (2.33) em $B := M_c[\mathrm{R}^k] \times L_2[\mathrm{R}^{k\times n}]$, e Φ é portanto bem definida. Note que definimos Φ de tal forma que (Y_t, Z_t) é uma solução de (2.26), (2.27) se, e só se, for um ponto fixo do mapa Φ. Vamos então mostrar que Φ é uma contração no espaço B, o qual muniremos com uma norma apropriada. Sejam $(U, V), (U', V') \in B$, $(Y, Z) = \Phi(U, V), (Y', Z') = \Phi(U', V')$, e $y = Y - Y', z = Z - Z'$. Precisamos calcular as normas de y e z em função das normas de $u = U - U'$ e $v = V - V'$. Perceba que

$$dy_t = d(Y_t - Y_t') = -(f(t, U_t, V_t) - f(t, U_t', V_t'))dt + (Z_t - Z_t')dW_t$$
$$= -(f(t, U_t, V_t) - f(t, U_t', V_t'))dt + z_t dW_t.$$

Estamos considerando, em B, uma métrica usual qualquer do produto. A ideia da demonstração é encontrar uma métrica equivalente em B, através da multiplicação dos processos originais por exponenciais, e em relação à qual o mapa Φ é uma contração. Começaremos modificando os processos dados com uma exponencial genérica e^{at}, para posteriormente escolher um a adequado. Como e^{at} é determinístico, $d\langle e^{at}, y\rangle_t = 0$. Além disso,

$$d\langle y\rangle_t = (z_t)^2 dt.$$

Assim, pela fórmula de Itô,

$$\begin{aligned} d(e^{at}|y_t|^2) &= e^{at}d(|y_t|^2) + ae^{at}|y_t|^2 dt = e^{at}(2y_t dy_t + (z_t)^2 dt) + ae^{at}|y_t|^2 dt \\ &= \{ae^{at}|y_t|^2 + e^{at}(z_t)^2 - 2e^{at}y_t(f(t,U_t,V_t) - f(t,U'_t,V'_t))\}dt + 2e^{at}y_t \cdot z_t dW_t. \end{aligned}$$

Dessa forma, reorganizando termos,

$$e^{at}(z_t)^2 dt - d(e^{at}|y_t|^2) = e^{at}(-a|y_t|^2 + 2y_t(f(t,U_t,V_t) - f(t,U'_t,V'_t)))dt - 2e^{at}y_t \cdot z_t dW_t, \tag{2.35}$$

e podemos integrar de t a T, levando em consideração que $y_T = Y_T - Y'_T = \xi - \xi = 0$:

$$e^{at}|y_t|^2 + \int_t^T e^{as}(z_s)^2 ds = \tag{2.36}$$

$$\int_t^T e^{as}(-a|y_s|^2 + 2y_s(f(s,U_s,V_s) - f(s,U'_s,V'_s)))ds - \int_t^T 2e^{as}y_s \cdot z_s dW_s. \tag{2.37}$$

Vamos tentar majorar parte do segundo membro dessa equação:

$$\begin{gathered} \int_t^T e^{as}2y_s(f(s,U_s,V_s) - f(s,U'_s,V'_s))ds \leqslant \\ \int_t^T 2e^{as}|y_s||f(s,U_s,V_s) - f(s,U'_s,V'_s)|ds \leqslant \\ \int_t^T 2Le^{as}|y_s|(|u_s| + |v_s|)ds. \end{gathered}$$

Substituindo em (2.36),

$$\begin{aligned} &e^{at}|y_t|^2 + \int_t^T e^{as}(z_s)^2 ds \leqslant \\ &\int_t^T -ae^{as}|y_s|^2 ds + \int_t^T 2Le^{as}|y_s|(|u_s| + |v_s|)ds - \int_t^T 2e^{as}y_s \cdot z_s dW_s. \end{aligned}$$

Aplicando a desigualdade $a^2 - 2ab\epsilon + \epsilon^2 b^2 = (a - \epsilon b)^2 \geqslant 0 \Rightarrow a^2/\epsilon + \epsilon b^2 \geqslant 2ab$ nos produtos $2|y_s||u_s|$ e $2|y_s||v_s|$ acima,

$$\begin{aligned} &e^{at}|y_t|^2 + \int_t^T e^{as}(z_s)^2 ds \leqslant \\ &\int_t^T -ae^{as}|y_s|^2 ds + \int_t^T Le^{as}(\epsilon|u_s|^2 + 2|y_s|^2/\epsilon + \epsilon|v_s|^2)ds - \int_t^T 2e^{as}y_s \cdot z_s dW_s \\ &= \int_t^T e^{as}|y_s|^2(2L/\epsilon - a)ds + \epsilon\int_t^T e^{as}(|u_s|^2 + |v_s|^2)ds - \int_t^T 2e^{as}y_s \cdot z_s dW_s. \end{aligned}$$

Note que, fixado um $\epsilon > 0$, a desigualdade acima vale para todo $a \in \mathrm{R}$. Em particular, podemos tomar $a = a(\epsilon) = 2L/\epsilon$ para obter

$$e^{at}|y_t|^2 + \int_t^T e^{as}(z_s)^2 ds \leqslant \epsilon \int_t^T e^{as}(|u_s|^2 + |v_s|^2)ds - \int_t^T 2e^{as}y_s \cdot z_s dW_s. \tag{2.38}$$

Denotanotando a integral $\int_0^T \epsilon e^{as}(|u_s|^2 + |v_s|^2)ds$ por $R[\epsilon]$, temos, com $a = a(\epsilon)$

$$\mathrm{E}[e^{at}|y_t|^2 + \int_t^T e^{as}(z_s)^2 ds] \leqslant \mathrm{E}[R[\epsilon]] - \mathrm{E}[\int_t^T 2e^{as}y_s \cdot z_s dW_s],$$

válida para todo $t \in [0, T]$. Acontece que, pelo lema 2.3.5, o processo $M_t = \int_0^t 2e^{as}y_s \cdot z_s dW_s$ é um martingale, de forma que a última esperança, que se escreve como $\mathrm{E}[M_T - M_t] = \mathrm{E}[M_0] - \mathrm{E}[M_0]$, é 0. Assim, podemos subtrair o primeiro termo do primeiro membro, que é positivo, da equação acima, sem alterar a desigualdade, e avaliar em $t = 0$:

$$\mathrm{E}[\int_0^T e^{as}|z_s|^2 dt] \leqslant R[\epsilon]. \tag{2.39}$$

Na inequação 2.38, podemos descartar o segundo termo do primeiro membro (também positivo) mantendo a desigualdade, usar a definição de $R[\epsilon]$ e colocar um módulo na última integral, obtendo

$$e^{at}|y_t|^2 \leqslant R[\epsilon] - 2\int_0^T e^{as}y_s \cdot z_s dW_s + 2|\int_0^t 2e^{as}y_s \cdot z_s dW_s|.$$

Tomando supremos e depois esperanças, e usando o fato de que a integral de Itô está definindo uma martingale,

$$\mathrm{E}[\sup_{t\in[0,T]} e^{at}|y_t|^2] \leqslant \mathrm{E}[R[\epsilon]] + 2\mathrm{E}[\sup_{t\in[0,T]} |\int_0^t e^{as}y_s \cdot z_s dW_s|].$$

Note que este último proceso é um martingale cuja variação quadrática é $\int_0^t e^{2as}|y_s|^2|z_s|^2 ds$. Usando então a desigualdade de Burkholder-Davis-Gundy, encontramos uma constante C (que não depende de nenhum dos parâmetros ϵ, a, etc; depende apenas de $p = 1$) tal que

$$\begin{aligned}
\mathrm{E}[\sup_{t\in[0,T]} e^{at}|y_t|^2] &\leqslant \mathrm{E}[R[\epsilon]] + C\mathrm{E}[(\int_0^T e^{2as}|y_s|^2|z_s|^2 ds)^{1/2}] \\
&\leqslant \mathrm{E}[R[\epsilon]] + \frac{1}{2}\mathrm{E}[2(\sup_{t\in[0,T]} e^{at/2}|y_s|)(\int_0^T C^2 e^{as}|z_s|^2 ds)^{1/2}] \\
&\leqslant \mathrm{E}[R[\epsilon]] + \frac{1}{2}\mathrm{E}[\sup_{t\in[0,T]} e^{at}|y_s|^2] + \frac{C^2}{2}\mathrm{E}[\int_0^T e^{as}|z_s|^2 ds].
\end{aligned}$$

Assim, multiplicando por 2 e rearranjando,

$$\mathrm{E}[\sup_{t\in[0,T]} e^{at}|y_t|^2] \leqslant 2\mathrm{E}[R[\epsilon]] + C^2\mathrm{E}[\int_0^T e^{as}|z_s|^2 ds]. \tag{2.40}$$

Somando (2.39) com (2.40) e usando (2.39) no segundo membro da desigualdade,

$$\mathrm{E}[\sup_{t\in[0,T]} e^{at}|y_t|^2 + \int_0^T e^{as}|z_s|^2 dt] \leqslant 3\mathrm{E}[R[\epsilon]] + C^2\mathrm{E}[\int_0^T e^{as}|z_s|^2] \leqslant (3+C^2)\mathrm{E}[R[\epsilon]]. \quad (2.41)$$

Isso nos sugere majorar $R[\epsilon]$. Então

$$R[\epsilon] = \int_0^T \epsilon e^{as}(|u_s|^2 + |v_s|^2)ds \leqslant \epsilon(T \sup_{t\in[0,T]} e^{as}|u_s|^2 ds + \int_0^T e^{as}|v_s|^2 ds).$$

Substituindo em (2.41),

$$\mathrm{E}[\sup_{t\in[0,T]} e^{at}|y_t|^2 + \int_0^T e^{as}|z_s|^2 dt] \leqslant \epsilon(3+C^2)(T+1)\mathrm{E}[\sup_{t\in[0,T]} e^{as}|u_s|^2 ds + \int_0^T e^{as}|v_s|^2 ds].$$

Como C é uma constante conhecida a priori e T é dado no início do problema, podemos escolher agora um ϵ tal que $\epsilon(3+C^2)(T+1) < 1$. De posse de tal ϵ, escolhemos $a = a(\epsilon)$ como indicado anteriormente. E, dessa forma, essa última desigulade se traduz para

$$\mathrm{E}[\sup_{t\in[0,T]} e^{at}|y_t|^2 + \int_0^T e^{as}|z_s|^2 dt] < \mathrm{E}[\sup_{t\in[0,T]} e^{as}|u_s|^2 ds + \int_0^T e^{as}|v_s|^2 ds],$$

ou seja, o espaço $B = M_c[\mathrm{R}^k] \times L_2[\mathrm{R}^{k\times n}]$ com a norma

$$|(Y,Z)|_a = E[\sup_{t\in[0,T]} e^{at}|Y_t|^2 + \int_0^T e^{as}|Z_s|^2 dt],$$

equivalente à norma do usual do produto cartesiano (já que temos normas pré-fixadas em M_c^2 e L^2), tem Φ como uma contração (basta lembrar da definição de y, z, u e v). Sendo B completo com essa norma, usamos o teorema do ponto fixo de Banach para obter **um único** ponto $(Y,Z) \in B$ tal que $\Phi(Y,Z) = (Y,Z)$. Ou seja, obtemos, como anunciado, uma única solução pro BSDE (2.26), (2.27). O resultado segue. □

Encerraremos a seção com um corolário óbvio, que nos diz para não nos preocuparmos muito com o espaço em que o processo Y_t, solução do BSDE, se encontra.

Corolário 2.3.9. *Sob as hipóteses usuais, o BSDE* (2.26), (2.27) *possui uma única solução tal que* $Z_t \in L^2[\mathrm{R}^{k\times n}]$

Demonstração. Já sabemos, pelo teorema 2.3.8, que o BSDE possui uma única solução, (Y_t, Z_t) em $M_c[\mathrm{R}^k]\times L_2[\mathrm{R}^{k\times n}]$. Seja agora (Y_t', Z_t') uma outra solução, tal que $Z_t' \in L^2[\mathrm{R}^{k\times n}]$. Como temos as condições usuais, o gerador f satisfaz

$$|f(t,y,z)| \leqslant |f(t,y,z) - f(t,0,0)| + |f(t,0,0)| \leqslant L(|y| + |z|) + |f(t,0,0)|,$$

onde o processo $G_t = f(t,0,0)$ está em $L^2(\Omega \times [0,T])$. Logo, pela proposição 2.3.4, $Y' \in M_c^2[\mathrm{R}^k]$, e temos $(Y', Z') = (Y,Z)$, por unicidade. □

2.4 Revisitando a EDP semilinear

Retornamos, nesta seção, a um desenvolvimento seguindo (PERKOWSKI, 2011, 2.3). Agora que entendemos melhor o BSDE (2.24), (2.25), podemos voltar à EDP semilinear

$$\partial_t v(t,x) + \mathcal{L}_t v(t,x) + f(t,x,v(t,x),D_x v(t,x)\sigma(t,x)) = 0, \qquad (t,x) \in [0,T], \tag{2.42}$$

que consideraremos com a condição terminal

$$v(T,x) = \phi(T). \tag{2.43}$$

Lembremos que o problema a ser resolvido, tendo em vista uma representação estocástica da solução dessa EDP, é do tipo

$$\begin{aligned} &dX_s^{t,x} = b(s,X_s^{t,x})ds + \sigma(X_s^{t,x})dW_s, \text{ para } s > t \\ &X_s = x, \text{ para } s \leqslant t \\ &dY_s^{t,x} = -f(s,X_s^{t,x},Y_s^{t,x},Z_s^{t,x}) + Z_s^{t,x}dW_s, \text{ para } 0 \leqslant s \leqslant T, \\ &Y_T^{t,x} = \phi(X_T^{t,x}), \end{aligned} \tag{2.44}$$

onde $b : [0,T] \times \mathrm{R}^d \to \mathrm{R}^d, \sigma : [0,T] \times R^d \to \mathrm{R}^{d\times n}, f : [0,T] \times \mathrm{R}^d \times \mathrm{R} \times \mathrm{R}^n \to \mathrm{R}$ são funções mensuráveis que satisfazem qtp as condições

$$\begin{aligned} &|b(t,x)| + |\sigma(t,x)| \leqslant C(1+|x|), \\ &|b(t,x) - b(t,x')| + |\sigma(t,x) - \sigma(t,x')| \leqslant C|x-y|, \\ &|f(t,x,y,z) - f(t,x,y',z')| \leqslant C(|y-y'| + |z-z'|), \\ &|f(t,x,0,0)| + |\phi(x)| \leqslant C(1+|x|^p), \end{aligned} \tag{2.45}$$

para todo x, y, y', z, z' e para algum $p \geqslant 1/2$. Em primeiro lugar, chamamos o tipo de sistema (2.44) de **equação diferencial estocástica forward-backward, FBSDE**, já que X é dado com condição inicial e Y com condição final. Dizemos ainda que o sistema é desacoplado, por X não depender de Y ou Z. Agora, sob as condições (2.45), temos existência e unicidade de um processo $X^{t,x}$ satisfazendo a SDE. Ainda $\mathcal{L}_t$ é o gerador infinitesimal do processo $X^{t,x}$. Essa solução satisfaz também a condição

$$\mathrm{E}[\sup_{s\geqslant t} |X_s|^{2p}] \leqslant K(1+|x|^{2p}). \tag{2.46}$$

Ainda, cosiderando o processo

$$g(\omega,t,y,z) = f(t,X^{t,x}(\omega),y,z),$$

vemos que

$$|g(\omega,t,y,z) - g(\omega,t,y',z')| = |f(t,X^{t,x}(\omega),y,z) - f(t,X^{t,x}(\omega),y',z')| \leqslant C(|y-y'|+|z-z'|),$$

satisfazendo a condição Lipscithz, e

$$|g(\omega, t, 0, 0)| = |f(t, X^{t,x}(\omega), 0, 0) \leqslant C(1 + |X^{t,x}(\omega)|^p),$$

que claramente pertence a $L^2(\Omega \times [0, T])$, por (2.46). Além disso $|\Phi(X_T)| \leqslant C(1 + |X_T|^p)$, e

$$\mathrm{E}[|X_T|^{2p}] \leqslant \sup_{t \leqslant s \leqslant T} \mathrm{E}[|X_s|^{2p}] \leqslant \mathrm{E}[\sup_{s \geqslant t} |X_s|^{2p}],$$

que é finita também por (2.46). Com isso, a condição terminal do BSDE que compõe o problema diferencial estocástico acima é quadrado integrável e temos existência e unicidade de solução, pela seção anterior.

Teorema 2.4.1. *Seja* $v \in C^{1,2}([0, T] \times \mathrm{R}^d)$ *uma solução de* (2.42), (2.43) *que satisfaz*

$$|D_x v(s, x)||\sigma(s, x)| \leqslant C(1 + |x|^k) \qquad \forall (s, x) \in [0, T] \times \mathrm{R}^d, \tag{2.47}$$

para algum $k > 1/2$. *Se* $(X^{t,x}_s, Y^{t,x}_s, Z^{t,x}_s)$ *é solução do problema estocástico* (2.44), *então* $v(t, x) = Y^{t,x}_t$ P-*qtp.*

Demonstração. Pela fórmula de Itô,

$$\begin{aligned} dv(s, X^{t,x}_s) &= (\partial_t v(s, X^{t,x}_s) + \mathcal{L}_s v(s, X^{t,x}_s))ds + D_x v(s, X^{t,x}_s)\sigma(s, X^{t,x}_s)dW_s = \\ -f(s, X^{t,x}_s&, v(s, X^{t,x}_s), D_x v(s, X^{t,x}_s)\sigma(s, X^{t,x}_s))ds + D_x v(s, X^{t,x}_s)\sigma(s, X^{t,x}_s)dW_s. \end{aligned} \tag{2.48}$$

Além disso,

$$v(T, X^{t,x}_T) = \phi(X^{t,x}_T). \tag{2.49}$$

Vemos então que o par de processos $(v(s, X^{t,x}_s), D_x v(s, X^{t,x}_s)\sigma(s, X^{t,x}_s))$ resolve o mesmo BSDE que $(Y^{t,x}_s, Z^{t,x}_s)$. Agora, note que

$$\mathrm{E}[\int_t^T |X^{t,x}_s|^{2k} ds] \leqslant (T - t)\mathrm{E}[\sup_{t \leqslant s \leqslant T} |X^{t,x}_s|^{2k}] \leqslant K(T - t)(1 + |x|^{2k}), \tag{2.50}$$

de forma que, por (2.47), obtemos que o processo $D_x v(s, X^{t,x}_s)\sigma(s, X^{t,x}_s)$ pertence a $L^2([t, T] \times \Omega)$. Assim, pelo Corolário 2.3.9, concluímos que $Y^{t,x}_s = v(s, X^{t,x}_s)$. A igualdade, contudo, se dá quando consideramos os processos em M^2_c, de forma que temos $\mathrm{E}[\sup_{t \leqslant s \leqslant T} |Y^{t,x}_s - v(t, X^{t,x}_s)|^2] = 0$, o que implica $\mathrm{E}[|Y^{t,x}_t - v(t, X^{t,x}_t)|^2] = \mathrm{E}[|Y^{t,x}_t - v(t, x)|^2] = 0$. Assim, $Y^{t,x}_t = v(t, x)$ P-qtp, e nosso resultado segue. □

Conseguimos, com esse teorema, uma representação estocástica para um tipo de equações diferenciais parciais semilineares. Na referência (PERKOWSKI, 2011), o mesmo teorema é demonstrado, mas lá pede-se uma hipótese adicional de limitação do crescimento da solução v em relação a x, uniforme em s. Conseguimos eliminar essa restrição justamente por conta do Corolário 2.3.9. Agora, notemos que o teorema nos garante dois corolários óbvios.

Corolário 2.4.2. *Se existir uma solução para* (2.42), (2.43) *que satisfaz* (2.47), *ela é única.*

Demonstração. De fato, se existirem duas soluções v_1 e v_2, para cada (t, x) existe um subconjundo de probabilidade 1 tal que $v_1(t, x) = Y_t^{t,x} = v_2(t, x)$. □

Corolário 2.4.3. *O processo* $Y_s^{t,x}$ *é tal que* $Y_t^{t,x}$ *é constante qtp.*

Em verdade, é possível mostrar que a variável aleatória $Y_t^{t,x}$ é de fato determinística.

2.5 Soluções de BSDEs

Apesar de terem sido introduzidas a fim de se representar estocasticamente soluções de EDPs, as BDSEs possuem razão própria de ser, sendo amplamente utilizadas em modelos financeiros e no estudo de problemas de controle estocástico, por exemplo. Dessa forma, nascem duas observações: (1) estudar formas de solução de BDSEs têm sua motivação intrínseca; e (2) partir da abordagem inversa do capítulo anterior e transformar num problema de EDPs o problema de solucionar BDSEs é perfeitamente razoável. Começamos então com um corolário direto da demostração do Teorema 2.4.1, que enunciaremos também como teorema.

Teorema 2.5.1. *Seja o BSDE definido por* $dY_s = -f(t, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})ds + Z_s dW_s$, $Y_T = \phi(X_s^{t,x})$ *em* $[0, T]$, *como em 2.44. Suponha que a EDP semilinear* (2.42) *com condição final* (2.43) *possui uma solução* $v \in C^{1,2}([0, T] \times \mathrm{R}^d)$ *que satisfaz a restrição de crescimento* (2.47). *Então a solução* $(Y_s^{t,x}, Z_s^{t,x})$ *do BSDE é dada por*

$$
\begin{aligned}
Y_s^{t,x} &= v(s, X_s^{t,x}); \\
Z_s^{t,x} &= Dv(s, X_s^{t,x})\sigma(s, X_s^{t,x}).
\end{aligned}
\tag{2.51}
$$

Demonstração. Segue direto da unicidade de solução do BDSE. □

Quando estudamos equações diferenciais, esperamos uma teoria que nos permita resolver, no mínimo, algumas classes de problemas, geralmente os mais simples. Em particular, é comum encontrarmos fórmulas explícitas para soluções de equações diferenciais lineares. Esse é também o caso com BSDEs. Encontraremos então fórmulas para soluções de BSDEs lineares. Antes, porém, definiremos precisamente qual classe de problemas é esse. Tendo em vista aplicações futuras, trabalharemos com BSDEs que geram processos Y_t na reta e nos preocuparemos com a forma deste processo Y_t.

Definição 2.5.2. *Seja* $f : \Omega \times [0, T] \times \mathrm{R} \times \mathrm{R}^{n\times 1} \to \mathrm{R}$ *do tipo*

$$f(t, y, z) = a_t y + b_t \cdot z + c_t, \tag{2.52}$$

onde $a : \Omega \times [0,T] \to \mathrm{R}$, $b : \Omega \times [0,T] \to \mathrm{R}^{1\times n}$, *e* $c : \Omega \times [0,T] \to \mathrm{R}$ *são processos progressivamente mensuráveis. Seja* ξ $\mathcal{F}_T$*-mensurável. Então, a BSDE*

$$dY_t = -f(t, Y_t, Z_t)dt + Z_s dW_s,\ t \in [0,T], Y_T = \xi \tag{2.53}$$

*é dito **linear**.*

Teorema 2.5.3. *Considere o BSDE linear (2.53). Suponha que os processos* a_t *e* b_t *sejam limitados, que o processo* c_t *pertença a* $L^2[\mathrm{R}]$ *e que* ξ *esteja em* $L^2(\Omega \times [0,T])$*. Então, a BSDE admite solução única, com* Y_t *dado por*

$$Y_t = H_t^{-1}\mathrm{E}[\xi H_T + \int_t^T c_s H_s ds | \mathcal{F}_t], \tag{2.54}$$

onde o processo

$$H_t = \exp\{\int_0^t b_s \cdot dW_s - \frac{1}{2}\int_0^t |b_s|^2 ds + \int_0^t a_s ds\} \tag{2.55}$$

*é chamado de **deflator**.*

Antes da demonstração, precisamos de um lema técnico.

Lema 2.5.4. *Seja* G_t *o processo*

$$\exp\{\int_0^t b_s dW_s - \frac{1}{2}\int_0^t b_s^2 ds\}. \tag{2.56}$$

Então, G_t *é um martingale.*

Demonstração. Chamando $P_t = \int_0^t b_s dW_s - \frac{1}{2}\int_0^t b_s^2 ds$, temos $dP_t = b_t dW_t - \frac{1}{2} b_t^2 dt$ e $d\langle P \rangle_t = b_t^2 dt$. Assim

$$dG_t = \exp(P_t)dP_t + \frac{1}{2}\exp(P_t)d\langle P \rangle_t = G_t b_t dW_t.$$

Dessa forma, G_t pode ser escrito como uma integral estocástica (bem definida pois temos integrandos contínuos), e é, portanto, um martingale local. Note agora que o coeficiente $s(t,x) = xb_t$ da SDE satisfeita por G_t tem as seguintes propriedades:

- $|s(t,x) - s(t,y)| = |xb_t - yb_t| \leqslant |B||x-y|$;
- $|s(t,x)| \leqslant |B||x|$,

onde B é a constante que majora o processo b_t. Assim,

$$\mathrm{E}[\sup_{t\in[0,T]} |G_t|] \leqslant \mathrm{E}[\sup_{t\in[0,T]} |G_t|^2] \leqslant C(1 + |G_0|^2) = 2C < \infty. \tag{2.57}$$

Como G_t é um martingale local uniformemente integrável, é um martigale, e o resultado segue. $\square$

Voltemos para a demonstração do teorema 2.5.3.

Demonstração. Comecemos notando que como a_t e b_t são contínuos e limitados, as integrais que aparecem na definição de H_t são todas bem definidas. Chamando de $L_t := \int_0^t b_r \cdot dW_s - \frac{1}{2}\int_0^t |b_s|^2 ds + \int_0^t a_s ds$, temos $dL_t = b_t \cdot dW_t - \frac{1}{2} b_t^2 dt + a_t dt$ e $d\langle L\rangle_t = b_t^2 dt$. Aplicando a formula de Itô em (2.55),

$$dH_t = \exp(L_t) dL_t + \frac{1}{2}\exp(L_t) d\langle L\rangle_t = H_t(a_t dt + b_t \cdot dW_t). \tag{2.58}$$

Também, se A for a constante que majora a_t, notando que $|H_t|^2 = \exp\{2\int_0^t a_s ds\}|G_t|^2$, com a notação do lema anterior, temos $\sup_{t\in[0,T]} |H_t|^2 \leqslant \exp(2AT) \sup_{t\in[0,T]} |G_t|^2$. Este último é integrável por (2.57), donde concluimos que $H_t \in M_c^2(\mathrm{R})$. Agora, percebamos que o gerador f do BSDE satisfazas condições usuais, já que os processos a_t e b_t são majorados ($\Rightarrow$ Lispchitz) e c_t é quadrado integrável ($\Rightarrow$ f(t, 0, 0) quadrado integrável). Temos então garantida a existência de uma solução (Y, Z) para o BSDE, onde Y também pertence a $M_c^2(\mathrm{R})$. Um cálculo simples nos diz que $d\langle Y, H\rangle_t = Z_t \cdot b_t dt$. Logo

$$\begin{aligned}
&d(H_tY_t) = H_t dY_t + Y_t dH_t + d\langle Y, H\rangle_t = \\
&- H_t(a_tY_t + b_t \cdot Z_t + c_t)dt + H_tZ_t dW_t + Y_tH_t(a_t dt + b_t dW_t) + H_tZ_t \cdot b_t dt = \\
&- H_tc_t dt + H_t(Z_t + Y_tb_t)dW_t \Rightarrow \\
&H_tY_t + \int_0^t c_sH_s ds = Y_0 + \int_0^t H_s(Z_s + Y_sb_s)dW_s.
\end{aligned} \tag{2.59}$$

Este último processo é, então, um martingale local. Porém

$$\mathrm{E}[\sup_{t\in[0,T]} |H_tY_t|] \leqslant C\mathrm{E}[\sup_{t\in[0,T]} |H_t^2 + Y_t^2|] \leqslant C(\mathrm{E}[\sup_{t\in[0,T]} H_t^2] + \mathrm{E}[\sup_{t\in[0,T]} Y_t^2]) < \infty, \tag{2.60}$$

e

$$\begin{aligned}
&\mathrm{E}[\sup_{t\in[0,T]} |\int_0^t c_sH_s ds|] \leqslant \mathrm{E}[\sup_{t\in[0,T]} |\sup_{s\in[0,T]} H_s \int_0^t c_s ds|] \leqslant \\
&C\mathrm{E}[\sup_{t\in[0,T]} (\sup_{s\in[0,T]} H_s^2 + (\int_0^t c_s ds)^2)] \leqslant \\
&C\mathrm{E}[\sup_{t\in[0,T]} H_t^2] + C\mathrm{E}[\sup_{t\in[0,T]} t\int_0^t c_s^2 ds] \leqslant C\mathrm{E}[\sup_{t\in[0,T]} H_s^2] + CT\mathrm{E}[\int_0^T c_s^2 ds] < \infty.
\end{aligned} \tag{2.61}$$

Concluimos que o processo $H_tY_t + \int_0^t c_sH_s ds$ é um martingale. Portanto,

$$H_tY_t + \int_0^t c_sH_s ds = \mathrm{E}[H_TY_T + \int_0^T c_sH_s ds|\mathcal{F}_t] = E[H_T\xi + \int_0^T c_sH_s ds|\mathcal{F}_t].$$

Passando a integral do primeiro para o segundo membro e inserindo-a dentro da esperança condicional, já que é $\mathcal{F}_t$-mensurável, e multiplicando por H_t^{-1}, temos

$$Y_t = H_t^{-1}E[H_T\xi + \int_t^T c_s H_s ds|\mathcal{F}_t],$$

como queríamos. □

2.6 Propriedades

Nesta seção, vamos mostrar um teorema de comparação para o comportamento das soluções de BSDEs em função de seus geradores e condições terminais. Esse resultado e seus corolários aparecerão de maneira bem sutil na modelagem apresentada na seção 2.7. Seguiremos aqui o desenvolvimento encontrado em (PHAM, 2009, 6.2.3).

Teorema 2.6.1. *Sejam (f^1, ξ^1) e (f^2, ξ^2) dois pares de geradores e condições terminais para BSDEs que satisfazem as condições usuais. Sejam (Y^1, Z^1) e (Y^2, Z^2) as respectivas soluções (tais que $Z^i \in L^2$). Suponha que*

- $\xi^1 \leqslant \xi^2$ *qtp;*
- $f^1(t, Y_t^1, Z_t^1) \leqslant f^2(t, Y_t^1, Z_t^1)$, $\mathrm{P} \times dt$- *qtp.*

Então, $Y_t^1 \leqslant Y_t^2$ para todo $t \in [0, T]$, qtp.

Demonstração. Por simplicidade, e tendo em vista aplicações na próxima seção, consideraremos processos Y e Z em R. Defina $\Delta Y = Y^2 - Y^1$, $\Delta Z = Z^2 - Z^1$. Considere os processos

$$\Delta_t^y = \frac{f^2(t, Y_t^2, Z_t^2) - f^2(t, Y_t^1, Z_t^2)}{Y_t^2 - Y_t^2} 1_{\{Y_t^2 - Y_t^1 \neq 0\}} \tag{2.62}$$

$$\Delta_t^z = \frac{f^2(t, Y_t^1, Z_t^2) - f^2(t, Y_t^1, Z_t^1)}{Z_t^2 - Z_t^2} 1_{\{Z_t^2 - Z_t^1 \neq 0\}} \tag{2.63}$$

$$\Delta f_t = f^2(t, Y_t^1, Z_t^1) - f^1(t, Y_t^1, Z_t^1). \tag{2.64}$$

Notando que os denominadores dos processos acima são ΔY e ΔZ, que

$$d(\Delta Y_t) = -(f^2(t, Y_t^2, Z_t^2) - f^1(t, Y_t^1, Z_t^1))dt + \Delta Z dW_t,$$

e que os processos Δ_t^y e Δ_t^z se anulam quando $\Delta Y = 0$ ou $\Delta Z = 0$, respectivamente, vemos que $(\Delta Y, \Delta Z)$ satisfaz o BSDE linear

$$d(\Delta Y_t) = -(\Delta_t^y \Delta Y_t + \Delta_t^z \Delta Z_t + \Delta f_t)dt - \Delta Z_t dW_t,$$

além da condição terminal $Y_T = \xi^2 - \xi^1$. Perceba que

$$|\Delta_t^y| \leqslant \frac{|L\Delta Y_t|}{|\Delta Y_t|} 1_{\{Y_t^2 - Y_t^1 \neq 0\}} \leqslant L,$$

por f^2 ser Lipschitz. O mesmo vale para Δ_t^z. Ou seja, esse processos são limitados. Ainda,

$$|f^2(t, Y_t^1, Z_t^1)| \leqslant |f^2(t, 0, 0)| + L(|Y_t^1| + |Z_t^1|),$$

e

$$|f^1(t, Y_t^1, Z_t^1)| \leqslant |f^1(t, 0, 0)| + L(|Y_t^1| + Z_t^1),$$

isto é, Δf_t é limitado por uma soma de processos em L^2 e, portanto, está também em L^2. Assim, concluímos que $(\Delta Y, \Delta Z)$ satisfaz uma BSDE linear de condições apropriadas, de forma que, por 2.5.3, temos

$$\Delta Y_t = H_t^{-1} \mathrm{E}[H_T(\xi^2 - \xi^1) + \int_t^T H_s \Delta f_s ds | \mathcal{F}_t], \tag{2.65}$$

onde H_t é o deflator do BSDE, sempre positivo. Concluímos de $\xi^2 - \xi^1 \geqslant 0$ e $\Delta f_t \geqslant 0$ $\mathrm{P} \times dt$ qtp que a esperança condicional acima é não negativa, e o resultado segue. □

Temos agora alguns corolários da demonstração.

Corolário 2.6.2. *Nas hipóteses e notação do teorema anterior, se tivermos $Y_0^2 \leqslant Y_0^1$, entao $Y_t^1 = Y_t^2$, $\forall t \in [0, T]$.*

Demonstração. Segue claramente de (2.65), já que

$$0 \leqslant \Delta Y_t = H_t^{-1} \mathrm{E}[H_T(\xi^2 - \xi^1) + \int_t^T H_s \Delta f_s ds | \mathcal{F}_t] \leqslant$$
$$H_t^{-1} \mathrm{E}[H_T(\xi^2 - \xi^1) + \int_0^T H_s \Delta f_s ds | \mathcal{F}_t] = H_t^{-1} \Delta Y_0 = 0,$$

pois todos processos são positivos e a última igualdade segue da hipótese. □

Corolário 2.6.3. *Se $\mathrm{P}(\xi_1 < \xi_2) > 0$ ou $f^1(t, Y_t^1, Z_t^1) < f^2(t, Y_t^1, Z_t^2)$ num conjunto de medida positiva, entao $Y_0^1 < Y_0^2$.*

Demonstração. Também de (2.65), vemos que ΔY_0 será dado pela esperança de um função positiva num conjunto de probabilidade não nula. □

Um último corolário fácil, mas importante:

Corolário 2.6.4. *Se o par (ξ, f) satisfaz $\xi \geqslant 0$ qtp e $f(t, 0, 0) \geqslant 0$ qtp, então, $Y_t \geqslant 0$ para todo t qtp. Ainda, se $\mathrm{P}[\xi > 0] > 0$ ou $f(t, 0, 0) > 0$ qtp, então $Y_0 > 0$.*

2.7 O modelo de Black-Scholes

Nesta seção, demonstraremos a fórmula de Black-Scholes para a precificação de derivativos, que será obtida através do maquinário dos BSDEs. Apesar de esta não ser a forma original que o problema foi abordado, e nem a única (de um viés probabilístico, podemos desenvolver a fórmula para preços por teoria de Girsanov, por exemplo, ver (LAMBERTON; LAPEYRE, 1996, capítulo 4)), os BSDEs promovem uma forma simples, limpa e natural de entender a questão de precificação, e nos abrem margem para ver de que forma podemos usá-los para resolver questões financeiras. Vamos descrever o problema para o caso discreto. Resolvê-lo-emos, contudo, num contexto contínuo. Consideremos um espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$, onde são definidas uma sequência de variáveis aleatórias i.i.d., (C_n), tais que $\mathrm{P}(C_i = 1) = \mathrm{P}(C_i = -1) = 1/2$, e tomemos a filtração $\mathcal{F}_n$ gerada pela sequência; $\mathcal{F}_0 := \{\varnothing, \Omega\}$.

Considere um mercado onde são trocados, nos tempos $n = 0, 1, 2, 3, \ldots$, dois tipo de ativos, S e B, cujos valores são dados por

$$\begin{aligned} B_n &= (1 + r)^n, \\ S_n &= (1 + \mu)^n + X_n, \end{aligned} \tag{2.66}$$

e X_n é o passeio aleatório

$$X_n = C_1 + \ldots + C_n. \tag{2.67}$$

B_n é chamado de ativo sem risco (bond), S_n é chamado de ativo de risco (stock) (às vezes, genericamente, de ação), r é a taxa de juros sem risco e μ a taxa de juros com risco, ambas constantes positivas. Uma **estratégia** ou **portfólio** de investimentos é um processo previsível em R^2, isto é, Z_n é $\mathcal{F}_{n-1}$-adaptado, que consiste em unidades dos ativos sem risco e com risco em cada uma das suas componentes, respectivamente. A restrição de uma estratégia previsível se dá justamente por conta da natureza do processo. Ao fazer uma estratégia de investimentos, isto é, ao se escolher, no tempo n, a quantidade de cada um dos ativos presentes no portfólio, o agente tem em sua posse apenas as informações sobre o comportamente dos ativos obtidas até o instante $n - 1$, e, portanto, $Z_n \in \mathcal{F}_{n-1}$. Bem por sua definição, dada uma estratégia Z_n, o valor $V_n = Z_n \cdot (B_n, S_n)$ é chamado de **valor do portfólio**. Um **derivativo** ξ_n é um ativo comercializado no mercado cujo preço deriva de outro ativo. Quando temos um derivativo do ativo de risco, por exemplo, podemos escrever $\xi_n = f(S_n)$, onde $f : \mathrm{R} \to R$ é uma função mensurável. Um tipo comum de ativos são as chamadas opções de compra e venda, e tentaremos modelar seu preço.

Um portfólio é dito **autofinanciável** quando as mudanças de valor entre os instantes $n = 1, 2, 3, \ldots$ se dá apenas pela alteração de sua composição e dos preços dos ativos, ou seja, não há fluxo externo, de entrada ou saída, de dinheiro na carteira. Com isso, a única ferramenta do agente a fim de controlar o desenvolvimento de seu capital é alternar, em cada instante, a composição do seu portfólio sem alterar seu valor. Isso se expressa por

$$Z_n \cdot (B_n, S_n) = Z_{n+1} \cdot (B_n, S_n). \tag{2.68}$$

Multiplicando (2.68) por -1 e somando $V_{n+1} = Z_{n+1} \cdot (B_{n+1}, S_{n+1})$ nos seus dois membros, obtemos $V_{n+1} - V_n = Z_{n+1} \cdot (B_{n+1} - B_n, S_{n+1} - S_n)$, expressa por

$$\Delta V_{n+1} = Z_{n+1} \cdot (\Delta B_{n+1}, \Delta S_{n+1}). \tag{2.69}$$

Uma opção (europeia) de compra (do ativo S_n) com vencimento em $n = N$ e preço de vencimento K é a **opção**, mas não obrigação, de se comprar, no instante N, o ativo S pelo preço K. Denotaremos por ξ_n esse derivativo. Notemos que

1. Se no instante N o valor de S for $S_N > K$, então o detentor da opção de compra poderá exercer seu direito de comprar S por K reais e vender esse título, imediatamente, pelo seu preço de mercado S_N. Essa operação lhe proporcionará $S_N - K$ reais de lucro instantaneamente;

2. Se no instante N o valor de S for $S_N \leqslant K$, não haverá sentido em o detentor desse ativo comprar S por um valor maior que o de mercado, e portanto não obterá lucro nenhum.

A questão a ser considerada então é decidir um preço para o derivativo ξ. Ou seja, queremos saber qual seria o valor justo a se pagar, em $n = 0$, por um papel que nos da a opção de comprar a ação S por K reais no instante N. Pelas duas observações acima, de antemão já sabemos que o valor final deste derivativo, $\xi_N = f(S_N)$, onde $f = (x - K)^+$. A nossa ideia então é encontrar o menor valor x tal que se pode montar uma estratégia de investimentos começando com x e terminando, no momento $n = N$, com um portfólio de valor ξ_N.

Tendo entendido a questão da precificação da opção de compra, consideraremos agora um modelo contínuo. As transações econômicas no mercado podem ser realizadas em instantes $t \in [0, T]$. Dessa forma, o processo de transformar o preço discreto dos ativos em preços contínuos envolve considerar a "continuização" das taxas de juros e do passeio aleatório. Se quisermos encontrar o valor do ativo sem risco, por exemplo, no instante t, subdividimos o intervalos $[0, t]$ em n intervalos de tamanhos idênticos e ajustamos a taxa de juros original por unidade de tempo, r, para novos intervalos de comprimento t/n, obtendo a taxa rt/n. Assim, o valor do ativo B no instante t é aproximado por seu valor após n passagens de tempo com taxas de juros ajustadas. O limite em n nos dará o valor desejado. Temos então

$$B_t = \lim_{n \to \infty} (1 + rt/n)^n = e^{rt}. \tag{2.70}$$

Já o passeio aleatório, quando passado por um processo de "continuização", se transforma num múltiplo de movimento Browniano por σ, onde esse coeficiente é entendido como a taxa de volatilidade do mercado. Assim, o modelo de Black-Scholes que considera ativos de risco e sem risco a tempos contínuos supõe preços modelados por

$$\begin{aligned} dB_t &= rB_t dt, \\ dS_t &= \mu S_t dt + \sigma S_t dW_t. \end{aligned} \tag{2.71}$$

Heuristicamente, Black-Scholes nos diz que a variação em porcentagem do preço do ativo sem risco é constante e proporcional ao tempo considerado; a variação proporcional do preço do ativo com risco tem uma componente proporcional ao tempo considerado e uma outra componente proporcional a um "passeio aleatório continuizado". Nomeamos a constante $\theta = \mu - r/\sigma$ por **prêmio de risco**, já que ela mede quanto o ganho certo da ação supera o ganho da variável sem risco em relação à volatilidade do mercado. As estratégias de investimentos são agora processos contínuos e adaptados em R^2, Z_t, e o valor do portfólio continua sendo $V_t = Z_t \cdot (B_t, S_t)$. A condição de auto-financiamento de um portfólio (2.69) se expressa como

$$\begin{aligned} &dV_t = Z_t^1 dB_t + Z_t^2 dS_t = Z_t^1(rB_t dt) + Z_t^2(\mu S_t dt + \sigma S_t dW_t) = \\ &(rV_t + \theta\sigma Z_t^2 S_t)dt + \sigma Z_t^2 S_t dW_t = \\ &(rV_t + \theta U_t)dt + U_t dW_t, \end{aligned} \tag{2.72}$$

onde $U_t = \sigma p_t$, p_t a quantidade de dinheiro em ações. Ainda, queremos um portfólio com valor final $V_T = f(S_T) = (S_T - K)^+$. Logo, temos o seguinte BSDE a ser resolvido:

$$dV_t = -(-rV_t - \theta U_t)dt + U_t dW_t, V_T = f(S_T); \tag{2.73}$$

e queremos o valor V_0. Por restrições econômicas, é bem razoável esperar um processo U em L^2, já que a quantidade de dinheiro deve sempre ser limitada. Assim, vemos que o gerador (linear) da BSDE satisfaz as hipóteses do teorema 2.5.3. Além disso, a condição terminal também está em L^2. Podemos então calcular o deflator H_t:

$$H_t = \exp\{\int_0^t -\theta dW_s - \frac{1}{2}\int_0^t \theta^2 ds - \int_0^t r ds\} = \exp\{-\theta W_t - \frac{1}{2}\theta^2 t - rt\}; \tag{2.74}$$

e encontrar V_t:

$$V_t = H_t^{-1}\mathrm{E}[f(S_T)H_T|\mathcal{F}_t] = \mathrm{E}[f(S_T)H_{(T-t)}|\mathcal{F}_t]. \tag{2.75}$$

V_0 é dado por

$$V_0 = \mathrm{E}[f(S_T)H_T|\mathcal{F}_0] = \mathrm{E}[f(S_T)H_T]. \tag{2.76}$$

O processo S_t é um movimento Browniano geométrico, de forma que

$$S_T = S_0 \exp\{(\mu - \frac{\sigma^2}{2})T + \sigma W_T\} = \exp\{(\mu - \frac{\sigma^2}{2})T + \sigma W_T\}, \tag{2.77}$$

pois vamos considerar ambos ativos tendo valor inicial unitário. Então,

$$V_0 = \mathrm{E}[(S_T - K)^+ H_T] = \mathrm{E}[(S_T - K)H_T; S_T \geqslant K].$$

Ainda,

$$S_T \geqslant K \Leftrightarrow (\mu - \frac{\sigma^2}{2})T + \sigma W_T \geqslant \ln K \Leftrightarrow W_T \geqslant \frac{1}{\sigma}(\ln K + (\frac{\sigma^2}{2} - \mu)T) := \alpha.$$

Também,

$$S_T H_T = \exp\{(\mu - \frac{\sigma^2 - \theta^2}{2} - r)T\} \exp\{(\sigma - \theta)W_T\} := \exp\{\beta T\} \exp\{(\sigma - \theta)W_T\}$$

e

$$H_T = \exp(-\frac{1}{2}\theta^2 - r)T \exp\{-\theta W_T\} := \exp\{\gamma T\} \exp\{-\theta W_T\}.$$

Compilando,

$$\begin{aligned} &V_0 = e^{\beta T}\mathrm{E}[\exp\{(\sigma - \theta)W_T\}; W_T \geqslant \alpha] - K e^{\gamma T}\mathrm{E}[\exp\{-\theta W_T; W_T \geqslant \alpha\}] = \\ &\frac{e^{\beta T}}{\sqrt{2\pi T}} \int_\alpha^\infty e^{(\sigma-\theta)x} e^{\frac{-x^2}{2T}} dx - \frac{K e^{\gamma T}}{\sqrt{2\pi T}} \int_\alpha^\infty e^{\theta x} e^{\frac{-x^2}{2T}} dx, \end{aligned} \tag{2.78}$$

e obtemos assim uma fórmula para o valor justo para uma opção de compra de ação. Na verdade, para se ter certeza que este é o valor justo, precisamos de uma observação. Estamos supondo que o agente gostaria de receber, no instante T, o valor $(K - X_T)^+$, e que não aceitaria nada menos que isso. Dessa forma, o valor obtido V_0 é, por unicidade de solução de BSDEs, o **único** valor inicial que um portfólio deveria ter a fim de se gerar, em T, o valor $(K - X_T)^+$. A questão, contudo, é se não existe algum outro valor inicial de portfólio, menor que V_0, que gera, no tempo T, uma quantia **maior** de dinheiro que $(K - X_T)^+$. Ora, neste caso, jamais um agente racional pagaria V_0 pela opção europeia de compra. Essa questão, entretanto, é respondida negativamente pelo teorema de comparação de BSDEs: se considerássemos dois BSDEs, 1 e 2, ambos representando a grandeza de portfólios V^1 e V^2, e portanto com os mesmos geradores, teríamos duas condições finais distintas, ξ_1 representando a opção de compra europeia, e ξ_2 uma outra variável aleatória com fluxo de dinheiro objetivamente maior que ξ_1 ($\mathrm{P}(\xi_2 > \xi_1) > 0$). O teorema nos garantiria $V_0^2 > V_0^1$, o que nos faz concluir que, de fato, V_0 é um preço justo. Outras observações importantes obtidas a partir do princípio de comparação de BSDEs são que: 1. Notemos que se a opção de comprar tiver possibilidade real de resultar um valor positivo em T, então o investimento inicial V_0 tem que ser maior que 0, por 2.6.4, o que confirma a hipótese de que se é impossível obter um capital positivo, sem risco, a partir de nenhum capital. Chamamos isso de **ausência de arbitragem**. Também a partir do corolário 2.6.4, notamos que $V_t \geqslant 0$, o que nos diz que o portfólio nunca precisa "tomar empréstimos". Esse tipo de estratégia é chamada de **admissível**. O interessante é que supor estratégias admissíveis dentro de mercados livre de arbitragem são, comumente, condições necessárias e impostas para conseguirmos encontrar o preço de derivativos em muitos dos caminhos utilizados para se precificá-los. Contudo, o método das BSDEs não nos forçou a adotar essas hipóteses; pelo contrário, nos mostrou que elas aparecem espontaneamente e como consequência da modelagem. Terminamos essa seção dizendo que é possível, além do valor inicial Y_0 da opção de compra, obtermos a **estratégia** necessária para, a partir de Y_0, obter-se $(K - B_T)^+$ no instante T. Essa conta, simples, pode ser encontrada em (LAMBERTON; LAPEYRE, 1996, Capítulo 4).

3 BSDEs e Problemas de Controle Estocásticos

A fim de ilustrar como as BSDEs também são utilizados para se resolver problemas de controle estocástico, vamos demonstrar um teorema que relaciona os dois conceitos e apresentar um outro exemplo em finanças. Seguiremos precisamente a abordagem dada por (PHAM, 2009, Seção 6.4.2) para o teorema. Fixamos, como sempre, um espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$ onde um movimento Browniano n-dimensional W_t é definido, e consideramos em Ω a filtração Browniana $(\mathcal{F}_t)$.

Uma SDE controlada por α é uma família de processos $X = X^\alpha$ em R^d dados por

$$dX_t = b(X_t, \alpha_t)dt + \sigma(X_t, \alpha_t)dW_t, t \in [0, t] \tag{3.1}$$

onde cada α é um processo $\alpha : \Omega \times [0, T] \to A \subseteq \mathrm{R}^m$ progressivamente mensurável, com A fixado, que chamamos de **controle do processo** (3.1). Temos $b : \mathrm{R}^d \times A \to \mathrm{R}^d$, $\sigma : \mathrm{R}^d \times A \to \mathrm{R}^{d \times n}$ mensuráveis e pedimos, como sempre,

$$|b(x, a) - b(y, a)| + |\sigma(x, a) - \sigma(y, a)| \leqslant L|x - y|$$
$$|b(x, a)| + |\sigma(x, a)| \leqslant K|x|$$

para todo $a \in A$, $x, y \in \mathrm{R}^d$, onde $L, K > 0$ são constantes. Como todo problema de controle, nosso objetivo é encontrar um **controle** α que **maximize** o funcional

$$J(\alpha) = \mathrm{E}\Big[\int_0^T f(t, X_t, \alpha_t)dt + g(X_T)\Big],$$

onde X_t acima é a difusão que resolve (3.1) com o α dado, $f : [0, T] \times \mathrm{R}^d \times A \to \mathrm{R}$ é contínua em (t, x) para todo $a \in A$, além de mensurável, $g : \mathrm{R}^d \to \mathrm{R}$ é uma função côncava C^1, e ambas f e g satisfazem uma condição de crescimento quadrática em x (dessa forma a esperança sempre é finita). Em resumo, se $\mathcal{A}$ é o conjunto dos processos progressivamente mensuráveis em A, queremos um processo $\hat{\alpha}$ tal que

$$J(\hat{\alpha}) = \sup_{\alpha \in \mathcal{A}} J(\alpha).$$

Definimos o **Hamiltoniano** $\mathcal{H} : [0, T] \times \mathrm{R}^d \times A \times \mathrm{R}^d \times \mathrm{R}^{d \times n} \to \mathrm{R}$ por

$$\mathcal{H}(t, x, a, y, z) = b(x, a) \cdot y + \mathrm{tr}(\sigma(x, a)^T z) + f(t, x, a), \tag{3.2}$$

e assumimos que $\mathcal{H}$ é diferenciável em x. Para cada $\alpha \in \mathcal{A}$, considere a seguinte BSDE, chamada de **BSDE adjunta**:

$$-dY_t = D_x\mathcal{H}(t, X_t, \alpha_t, Y_t, Z_t)dt - Z_t dW_t, \qquad Y_T = D_x g(X_T). \tag{3.3}$$

Temos então a ligação entre BSDEs e problemas de controle:

Teorema 3.1.1. *Seja $\hat{\alpha}_t \in \mathcal{A}$ e $\hat{X}_t$ o processo controlado por $\hat{\alpha}_t$ associado. Suponha que para tais $\hat{\alpha}_t$ e $\hat{X}_t$ exista uma solução $(\hat{Y}_t, \hat{Z}_t)$ para a BSDE (3.3) associada que satisfaz*

$$\mathcal{H}(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t) = \max_{a \in A} \mathcal{H}(t, \hat{X}_t, a, \hat{Y}_t, \hat{Z}_t) \tag{3.4}$$

e

$$(x, a) \to \mathcal{H}(t, x, a, \hat{Y}_t, \hat{X}_t) \text{ é uma função côncava, } \forall t \in [0, T]. \tag{3.5}$$

*Então, $\hat{\alpha}_t$ é um **controle ótimo.***

Demonstração. Seja α um controle qualquer para o sistema, com X_t associado. Escrevemos

$$J(\hat{\alpha}) - J(\alpha) = \mathrm{E}\Big[\int_0^T f(t, \hat{X}_t, \hat{\alpha_t}) - f(t, X_t, \alpha)dt + g(\hat{X}_T) - g(X_T)\Big]. \tag{3.6}$$

Usando a concavidade de g e a fórmula de Itô para o produto,

$$\begin{aligned} \mathrm{E}[g(\hat{X}_T) - g(X_T)] \geqslant \mathrm{E}[(\hat{X}_T - X_T)D_x g(\hat{X}_T)] = \mathrm{E}[(\hat{X}_T - X_T)\hat{Y}_T] = \\ \mathrm{E}[\int_0^T (\hat{X}_t - X_t)d\hat{Y}_t + \int_0^T \hat{Y}_t(d\hat{X}_t - dX_t) + \int_0^T \mathrm{tr}[(\sigma(\hat{X}_t, \hat{\alpha}_t) - \sigma(X_t, \alpha_t))^T \hat{Z}_t]dt] = \\ \mathrm{E}[\int_0^T (\hat{X}_t - X_t)(-D_x\mathcal{H}(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t))dt] + \int_0^t \hat{Y}_t(b(\hat{X}_t, \hat{\alpha}_t) - b(X_t, \alpha_t))dt \\ +\mathrm{E}[\int_0^T \mathrm{tr}[(\sigma(\hat{X}_t, \hat{\alpha}_t) - \sigma(X_t, \alpha_t))^T \hat{Z}_t]dt]. \end{aligned} \tag{3.7}$$

Ainda, temos, por (3.2)

$$\begin{aligned} \mathrm{E}[\int_0^T f(t, \hat{X}_t, \hat{\alpha_t}) - f(t, X_t, \alpha)dt] = \mathrm{E}[\int_0^T \mathcal{H}(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t)) - \mathcal{H}(t, X_t, \alpha_t, \hat{Y}_t, \hat{Z}_t)dt] - \\ \mathrm{E}[\int_0^t \hat{Y}_t(b(\hat{X}_t, \hat{\alpha}_t) - b(X_t, \alpha_t))dt - \int_0^T \mathrm{tr}[(\sigma(\hat{X}_t, \hat{\alpha}_t) - \sigma(X_t, \alpha_t))^T \hat{Z}_t dt]. \end{aligned} \tag{3.8}$$

Substituindo (3.7) e (3.8) em (3.6), temos

$$\begin{aligned} J(\hat{\alpha}) - J(\alpha)) \geqslant \mathrm{E}[\int_0^T \{\mathcal{H}(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t) - \mathcal{H}(t, X_t, \alpha_t, \hat{Y}_t, \hat{Z}_t)\} \\ -\{(\hat{X}_t - X_t)D_x\mathcal{H}(t, \hat{X}_t, \hat{\alpha}_t, \hat{Y}_t, \hat{Z}_t)\}dt]. \end{aligned} \tag{3.9}$$

Agora, a concavidade de $\mathcal{H}$ em (x, a) e a propriedade (3.4) de $\hat{\alpha}$ fazem o integrando acima ser $\geqslant 0$, de onde concluímos que $J(\hat{\alpha}) - J(\alpha) \geqslant 0$, e $\hat{\alpha}$ é de fato um controle ótimo. □

Agora, vamos apresentar um exemplo de uma ideia de como o teorema 3.1.1 pode ser usado para se resolver um problema simples em finanças. Contudo, não faremos todas as contas nem daremos uma solução final. Apenas gostaríamos de mostrar de que sorte são as abordagens de problemas de controle estocástico via BSDEs.

Exemplo: Considere, como na seção 2.7, um mercado com dois ativos, B_t e S_t, cujos preços são dados por

$$dB_t = rB_t dt,\ B_0 = 1 \tag{3.10}$$

$$dS_t = \mu S_t dt + \sigma S_t dW_t,\ S_0 = 1. \tag{3.11}$$

Ou seja, temos um ativo de risco e um ativo sem risco no mercado. Suponha ainda que a taxa de juros a risco é maior que a taxa de juros sem risco, i.e., $\mu > r$. Considere agora um investidor que começa com uma quantia inicial x e que, através de uma estratégia **auto-financiada**, deseja maximizar seus lucros no instante T. Para tanto, ele investe, a cada instante $t \in [0, T]$, todo o capital que possui. Sua única questão então é como decidir qual porcentagem $\alpha(t)$ investir no ativo de risco S_t e qual investir no ativo sem risco B_t. Claro que suas decisões são condizentes com a quantidade de informação que possui a cada instante, de forma que ele deseja então encontrar um **processo adaptado** $\alpha : \Omega \times [0, T] \to [0, 1]$ que maximize sua riqueza X_t em T, dada por

$$dX_t = \frac{(1-\alpha_t)X_t}{B_t} dB_t + \frac{\alpha_t X_t}{S_t} dS_t,\ X_0 = x. \tag{3.12}$$

Note que $(1-\alpha_t)X_t$ é a quantidade de dinheiro dispendida para comprar ativos sem risco, de forma que $\dfrac{(1-\alpha_t)X_t}{B_t}$ é a quantidade desses ativos comprados em t, e a equação está explicada. Agora, (3.12) pode ser combinada com (3.10) e (3.11) e re-escrita como

$$dX_t = [(\mu - r)\alpha_t + r]X_t dt + \sigma \alpha_t X_t dW_t,\ X_0 = x. \tag{3.13}$$

O funcional que gostaríamos de maximizar é simplesmente

$$J[\alpha] = \mathrm{E}[X_T] = \mathrm{E}[\int_0^T f(t, X_t, \alpha_t)dt + g(X_T)], \tag{3.14}$$

com $f = 0$ e $g = \mathrm{Id}$. Assim, escrevendo $b(x, a) = [(\mu - r)a + r]x$ e $\Sigma(x, a) = \sigma a x$, a equação (3.13) se torna

$$dX_t = b(\alpha_t, X_t)dt + \Sigma(\alpha_t, X_t)dW_t,\ X_0 = x. \tag{3.15}$$

Podemos então escrever o hamiltoniano deste sistema,

$$\mathcal{H}(t, x, a, y, z) = b(x, a)y + \Sigma(a, x)z + f(t, x, a) = [(\mu - r)a + r]xy + \sigma a x z, \tag{3.16}$$

e a BSDE associada

$$\begin{aligned} dY_t = -D_x\mathcal{H}(t, X_t, \alpha_t, Y_t, Z_t)dt + Z_t dW_t;\ Y_T = D_x g(X_T) \Rightarrow \\ dY_t = -\{[(\mu - r)\alpha_t + r]Y_t + \sigma\alpha_t Z_t\}dt + Z_t dW_t;\ Y_T = 0. \end{aligned} \tag{3.17}$$

Precisamos estudar a quantia

$$\max_{a\in[0,1]} \mathcal{H}(t, X_t, a, Y_t, Z_t) = \max_{a\in[0,1]} [(\mu - r)a + r]X_tY_t + \sigma a X_t Z_t. \tag{3.18}$$

Note que

$$[(\mu - r)a + r]X_tY_t + \sigma a X_t Z_t = a[(\mu - r)X_tY_t + \sigma X_t Z_t] + rX_tY_t, \tag{3.19}$$

de forma que a expressão dentro do máximo em (3.18) é linear em a. Sendo assim, a expressão assume máximo em $a = 0$ ou $a = 1$, a depender se o coeficiente linear $(\mu - r)X_tY_t + \sigma X_t Z_t$ é positivo ou negativo. Pelo teorema 3.1.1, queremos um controle $\alpha(t)$ tal que (3.4) seja satisfeita. Pela observação acima, concluímos que um **controle ótimo** para a dinâmica da riqueza do investidor assume apenas os valores 0 ou 1, significando que, a depender do estado do mundo, ora o investidor aplica toda sua riqueza no ativo de risco ($\alpha(t) = 1$), ora ele aplica toda sua riqueza no ativo sem risco ($\alpha(t) = 0$). A questão então é definir quando há a troca do valor de controle de, digamos, 1 para 0. Por (3.19). Vamos estudar então o sinal de

$$(\mu - r)X_tY_t + \sigma X_t Z_t. \tag{3.20}$$

Comecemos com um lema:

Lema 3.1.2. *Seja a SDE linear, iniciada em t_0*

$$dX_t = (a(t)X_t + c(t))dt + (b(t)X_t + d(t))dW_t$$

Então

$$X_t = \Phi_{t,t_0}\left(X_{t_0} + \int_{t_0}^{t} \Phi_{s,t_0}^{-1}(c(s) - b(s)d(s))ds + \int_{t_0}^{t} \Phi_{s,t_0}^{-1} d(s)dW_s\right),$$

onde

$$\Phi_{t,t_0} = \exp\left(\int_{t_0}^{t}\left(a(s) - \frac{b^2(s)}{2}\right)ds + \int_{t_0}^{t} b(s)dW_s\right).$$

Para a SDE (3.13), os coeficientes $c(t)$ e $d(t)$, na terminologia do lema, são nulos, de maneira que X_t é dado por uma exponencial, sendo sempre positivo. Assim, em (3.20), podemos cancelar X_t e nos preocupar apenas com o sinal da expressão

$$(\mu - r)Y_t + \sigma Z_t. \tag{3.21}$$

Para isso, precisamos conhecer (Y_t, Z_t), solução de (3.17) associada ao controle ótimo α. Se supormos Y_t e Z_t contínuos qtp (o que de fato vem a ser o caso) e supormos que o controle se inicie, por exemplo, em 1 (lembremos que por α ser adaptado, $\alpha(0)$ tem que ser ω-constante), teremos que o sinal de (3.21) se mantém constante no intervalo $[0, \tau(\omega)]$, e portanto o controle $\alpha(t)$ vale 1 em $[0, \tau(\omega)]$ e precisamos apenas encontrar o tempo de parada τ. Note que a BSDE (3.17) fica

$$dY_t = -(\mu Y_t + \sigma Z_t)dt + Z_t W_t, Y_T = 0. \tag{3.22}$$

Precisamos resolver a equação acima para encontrar o sinal de (3.21). Conhecemos uma expressão para Y_t, por ser um BSDE linear, mas não conhecemos uma expressão para Z_t. A ideia então é supormos que Y_t é do tipo $Y_t = u(t, X_t)$, com X_t dado por

$$dX_t = \mu X_t dt + \sigma X_t dW_t,\ X_0 = x,\ t \in [0, \tau(\omega)], \tag{3.23}$$

onde apenas substituímos o controle $\alpha(t) = 1$ pata $t \in [0, \tau(\omega)]$ em (3.13). Aplicando a fórmula de Itô para Y dessa forma, temos

$$\begin{gathered} dY_t = du(t, X_t) = \partial_t u(t, X_t)dt + \partial_x u(t, X_t)dX_t + \frac{1}{2}\partial_x^2 u(t, X_t)d\langle X\rangle_t = \\ (\partial_t u(t, X_t) + \partial_x u(t, X_t)\mu X_t + \frac{1}{2}\partial_x^2 u(t, X_t)\sigma^2 X_t)dt + (\partial_x u(t, X_t)\sigma X_t)dW_t. \end{gathered} \tag{3.24}$$

Comparando (3.24) com (3.22), vemos que gostaríamos que $u(t, x)$ também satisfizesse

$$\sigma X_t \partial_x u(t, X_t) = Z_t. \tag{3.25}$$

Dessa forma, se queremos uma solução (Y_t, Z_t) para (3.22) do tipo $Y_t = u(t, X_t)$, então precisamos de uma função $u(t, x)$ (suficientemente suave) que satisfaça

$$\begin{gathered} \partial_t u(t, x) + \mu x \partial_x(t, x) + \frac{1}{2}\sigma^2 x \partial_x^2 u(t, x) = \\ -(\mu u(t, x) + \sigma^2 x \partial_x u(t, x)), \qquad u(T, x) = 0, \end{gathered} \tag{3.26}$$

onde novamente comparamos (3.22) com (3.24) e enxergamos $Y_t = u(t, X_t)$ e Z_t como em (3.25). Assim, resolvendo o problema de valor terminal (3.26), obtemos uma função $u(t, x)$ tal que $u(t, X(t)) = Y_t$; $\sigma X_t \partial_x u(t, X_t) = Z_t$. Lembrando que o lema 3.1.2 nos dá uma solução exata para X_t, conseguimos (pelo menos em tese e resolvendo (3.26)) uma expressão exata para Y_t e Z_t, no intervalo $[0, \tau(\omega)]$, o que nos permite analisar o sinal de (3.21) e, finalmente, determinar o tempo de parada τ, que determina a mudança do controle α de 1 para 0. Lógico que após a determinação de τ, seguimos um processo parecido no intervalo $(\tau(\omega), \tau_1(\omega)]$ para determinar o intervalo exato, para cada estado de mundo, em que o controle α valerá 0, até ser trocado novamente para 1. E isso nos permite, a menos de contas, entender o problema financeiro de se determinar a melhor estratégia (α_t) de investimentos a fim de se maximizar o valor final da riqueza X_T.

Terminamos esse exemplo com um comentário sobre o método de se encontrar a solução da BSDE (3.22) através da suposição de solução do tipo $Y_t = u(t, X_t)$, transformando o problema da BSDE em um problema de EDP. Esse método é amplamente utilizado em teoria de BSDEs e pode ser mais bem compreendido através de um esquema chamado “esquema dos quatro passos”, explicado de forma cuidadosa em (HU; YONG, 2000), que mostra condições suficientes e necessárias sob as quais o método é válido, para casos bem gerais, inclusive onde os coeficientes das equações que definem X_t e (Y_t, Z_t) não são suaves.

4 Considerações Finais

Nestas considerações finais, gostaria de deixar três questionamentos, provavelmente ingênuos, que surgiram durante o estudo, a respeito de generalizações. São eles:

1. A característica do movimento Browniano W_t que confere ao cálculo em relação ao integrador dW_t propriedades totalmente distintas do cálculo usual é o fato de que, em certo sentido, $(dW_t)^2 = dt$. E esse tipo de diferencial tem em si todo um poder de criar um cálculo particular, com propriedades únicas (o que acaba permitindo a representação de EDPs, por exemplo), justamente porque vai de encontro à definição (ou tentativa de definição) clássica de diferenciais como grandezas dx diferentes de 0 tais que $(dx)^2 = 0$. E, embora não tenhamos exibido explicitamente a demonstração da fórmula de Itô, responsável pelas peculiaridades do cálculo de Itô "contra" o integrador dW_t, ela nasce da seguinte observação. Se f é de classe C^2, então, pela fórmula de Taylor,

$$f(W_t) = f(0) + f'(0)(\Delta W_t) + \frac{1}{2}f''(0)(\Delta W_t)^2 + o(\Delta W_t^3). \tag{4.1}$$

Dessa forma, tomando limites, e todos esses passos podem ser justificados de forma rigorosa,

$$df(W_t) = f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt.$$

A questão que surge então é sobre a existência de processos H_t tais que $(dH_t)^3 = dt$, ou, mais geralmente, processos H^n tais que $(dH_t^n)^n = dt$. E, se eles existem, como seria o cálculo relacionado, já que (4.1) teria uma expressão distinta.

2. Os semimartingales X_t foram definidos como processos da forma $M_t + B_t$, onde M_t é um martingale local e B_t um processo de variação limitada. Acontece que é possível desenvolver uma integral de Itô para processos progressivamente mensuráveis f (com algumas condições de integrabilidade) "contra" o integrador dM_t da mesma forma que desenvolvemos para o integrador dW_t. Para isso, definimos a integral para processos simples $\phi_t(\omega) = \sum_{i=0}^{n-1} \xi_i(\omega)1_{(t_i,t_{i+1}]}(t)$ em relação ao integrador dM_t como em (1.8):

$$(\int_0^T \phi_t dM_t)(\omega) := \sum_{i=0}^{n-1} \phi_i(\omega)(M_{t_{i+1}}(\omega) - M_{t_i}(\omega)). \tag{4.2}$$

Ainda, vale um análogo da isometria de Itô (sob condições de integrabilidade do processo),

$$\mathrm{E}[\int_0^T \phi^2 d\langle M\rangle_t] = \mathrm{E}[(\int_0^T \phi dM_t)^2],$$

de forma que conseguimos estender a integral para o fecho dos processos simples, e desenvolvemos uma teoria de integração estocástica de maneira inteiramente correlata à apresentada neste trabalho. Mais ainda, o processo B_t tem variação finita, de forma que ele gera uma medida de Lebesgue-Stieltjes na reta para quase todos ω, e podemos então definir a integral estocástica em relação ao integrador dX_t (novamente, sob condições de integrabilidade para o integrando):

$$\int_0^T f(\omega,t)dX_t(\omega) := \int_0^T f(\omega,t)dM_t(\omega) + \int_0^T f(\omega,t)dB_t(\omega).$$

Tendo observado isso, é interessante pontuar que há uma definição alternativa para semimartingales como "processos que geram integradores". Mais especificamente, diz-se que um processo X é um semimartingale se **há** o operador I_X que leva processos previsíveis limitados em v.a.'s quadrado integráveis de forma que: (i) o operador é linear e concorda com a expressão (4.2) para processos simples; (ii) se ξ_n é uma família uniformemente limitada de processos previsíveis tendendo a um limite ξ, então $I_X(\xi_n)$ tende a $I_X(\xi)$ em probabilidade. O **teorema de Bichteler-Dellacherie** fornece uma equivalência entre as duas definições. Toda essa excursão sobre semimartingales é feita de forma completa em (LOWTHER, 2011a) (na verdade, em toda a parte do blog sob a aba "The General Theory of Semimartingales"). A questão que surge então é: será possível desenvolver uma teoria de integração estocástica em relação a integradores que não sejam semimartingales? Qual a classe mais abundante de processos que serve como integradores, e como seria o cálculo relacionado?

3. O teorema de existência e unicidade de soluções para BSDEs se vale fortemente do teorema de representação de martingales A.3.8. Surge então a questão: se M_t é um processo estocástico satisfazendo determinado conjunto de propriedades a se determinar (ser martingale, ser càdlàg, etc) há algum teorema análogo à representação de martingales, ou seja, se N_t for um martingale em relação à filtração completada gerada por M_t, é possível escrever N_t como uma integral "contra" dM_t? Se sim, como poderíamos desenvolver a teoria de BSDEs relacionada?

Referências

BARTLE, R. *The elements of integration and Lebesgue measure*. Wiley, 1995. (Wiley classics library). ISBN 9780471042228. Disponível em: <https://books.google.com.br/books?id=8oNGAAAAYAAJ>. Citado na página 97.

BASS, R. *Stochastic Processes*. Cambridge University Press, 2011. (Cambridge Series in Statistical and Probabilistic Mathematics). ISBN 9781139501477. Disponível em: <https://books.google.com.br/books?id=Ll0T7PIkcKMC>. Citado 5 vezes nas páginas 25, 34, 35, 36 e 37.

BOULEAU, N. *Probabilités de l'ingénieur: Variables aléatoires et simulation*. Hermann, Éd. des Sciences et des Arts, 1986. (Actualités scientifiques et industrielles). ISBN 9782705614188. Disponível em: <https://books.google.com.br/books?id=nvZQAAAAMAAJ>. Citado na página 15.

CATUOGNO, P. J. *Notas de aula sobre Integração Estocástica*. 2013. Citado 7 vezes nas páginas 24, 25, 27, 36, 38, 40 e 42.

______. *Notas de aula sobre martingales*. 2015. Citado na página 23.

DOOB, J. L. What is a martingale? *The American Mathematical Monthly*, Mathematical Association of America, v. 78, n. 5, p. 451–463, 1971. ISSN 00029890, 19300972. Disponível em: <http://www.jstor.org/stable/2317751>. Citado na página 16.

EVANS, L. *An Introduction to Stochastic Differential Equations*. American Mathematical Society, 2012. ISBN 9781470410544. Disponível em: <https://books.google.com.br/books?id=VYoCAQAAQBAJ>. Citado na página 45.

FOLLAND, G. *Real Analysis: Modern Techniques and Their Applications*. Wiley, 2013. (Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts). ISBN 9781118626399. Disponível em: <https://books.google.com.br/books?id=wI4fAwAAQBAJ>. Citado na página 24.

FRIEDMAN, A. *Partial differential equations of parabolic type*. Prentice-Hall, 1964. Disponível em: <https://books.google.com.br/books?id=XN0pAQAAMAAJ>. Citado 2 vezes nas páginas 54 e 57.

HERVéS-BELOSO, C.; MONTEIRO, P. K. *Information and σ-algebras*. 2012. Citado 2 vezes nas páginas 13 e 15.

HU, Y. *Backward Stochastic Differential Equations and Applications in Finance*. 2013. Citado na página 62.

HU, Y.; YONG, J. Forward–backward stochastic differential equations with nonsmooth coefficients. *Stochastic Processes and their Applications*, v. 87, n. 1, p. 93 – 106, 2000. ISSN 0304-4149. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0304414999001064>. Citado na página 90.

KARATZAS, I.; SHREVE, S. *Brownian motion and stochastic calculus*. Springer-Verlag, 1988. (Graduate texts in mathematics). ISBN 9783540965350. Disponível em: <https://books.google.com.br/books?id=uxnvAAAAMAAJ>. Citado 3 vezes nas páginas 49, 54 e 57.

KAROUI, N. E.; PENG, S.; QUENEZ, M. C. Backward stochastic differential equations in finance. *Mathematical Finance*, Blackwell Publishers Inc., v. 7, n. 1, p. 1–71, 1997. ISSN 1467-9965. Disponível em: <http://dx.doi.org/10.1111/1467-9965.00022>. Citado 2 vezes nas páginas 12 e 46.

KUO, H. *Introduction to Stochastic Integration*. Springer New York, 2005. (Universitext). ISBN 9780387287201. Disponível em: <https://books.google.com.br/books?id=VEAxuzpCvj0C>. Citado 3 vezes nas páginas 18, 20 e 45.

LAMBERTON, D.; LAPEYRE, B. *Introduction to Stochastic Calculus Applied to Finance, Second Edition*. Taylor & Francis, 1996. (Chapman & Hall/CRC Financial Mathematics Series). ISBN 9780412718007. Disponível em: <https://books.google.com.br/books?id=61zI_o-pIkEC>. Citado 2 vezes nas páginas 80 e 84.

LOWTHER, G. *Optional Sampling*. 2009. <https://almostsure.wordpress.com/2009/12/20/optional-sampling/>. Accessed: 2018-02-18. Citado na página 99.

______. *The Burkholder-Davis-Gundy Inequality*. 2010. <https://almostsure.wordpress.com/2010/04/06/the-burkholder-davis-gundy-inequality/>. Accessed: 2018-02-18. Citado na página 100.

______. *The Bichteler-Dellacherie Theorem*. 2011. <https://almostsure.wordpress.com/2011/03/>. Accessed: 2018-02-18. Citado na página 92.

______. *The Doob-Meyer Decomposition*. 2011. <https://almostsure.wordpress.com/2009/12/22/class-d-processes/>. Accessed: 2018-02-18. Citado 2 vezes nas páginas 34 e 36.

MARTIN, S. L.; MARQUES, M. S. d. F. *Cálculo Estocástico*. SOCIEDADE BRASILEIRA DE MATEMATICA, 1991. ISBN 85-244-0059-5. Disponível em: <https://impa.br/wp-content/uploads/2017/04/18_CBM_91_07.pdf>. Citado na página 98.

MUNKRES, J. *Analysis On Manifolds*. Avalon Publishing, 1997. (Advanced Books Classics). ISBN 9780813345482. Disponível em: <https://books.google.com.br/books?id=tGT6K6HdFfwC>. Citado na página 55.

ØKSENDAL, B. *Stochastic Differential Equations: An Introduction with Applications*. Springer Berlin Heidelberg, 2010. (Universitext). ISBN 9783642143946. Disponível em: <https://books.google.com.br/books?id=EQZEAAAAQBAJ>. Citado na página 100.

PARDOUX, E.; PENG, S. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, v. 14, n. 1, p. 55 – 61, 1990. ISSN 0167-6911. Disponível em: <http://www.sciencedirect.com/science/article/pii/0167691190900826>. Citado na página 67.

PERKOWSKI, N. *Backward Stochastic Differential Equations: an Introduction*. 2011. Citado 4 vezes nas páginas 46, 60, 73 e 74.

PHAM, H. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer Berlin Heidelberg, 2009. (Stochastic Modelling and Applied Probability). ISBN 9783540895008. Disponível em: <https://books.google.com.br/books?id=xBsagiBp1SYC>. Citado 3 vezes nas páginas 12, 78 e 85.

STEELE, J. *Stochastic Calculus and Financial Applications*. Springer, 2001. (Applications of mathematics : stochastic modelling and applied probability). ISBN 9780387950167. Disponível em: <https://books.google.com.br/books?id=H06xzeRQgV4C>. Citado 7 vezes nas páginas 20, 27, 28, 29, 32, 59 e 99.

TESCHL, G. *Ordinary Differential Equations and Dynamical Systems*. American Mathematical Soc., 2012. (Graduate studies in mathematics). ISBN 9780821891056. Disponível em: <https://books.google.com.br/books?id=FSObYfuWceMC>. Citado na página 65.

Zhou, Y.; Cai, W. Numerical Solution of the Robin Problem of Laplace Equations with a Feynman-Kac Formula and Reflecting Brownian Motions. *ArXiv e-prints*, jun. 2015. Citado na página 48.

Apêndices

APÊNDICE A – Miscelânea de Resultados

A.1 Teoria da Medida

Vamos fixar um espaço de medida $(X, \mathcal{F}, \mu)$.

Teorema A.1.1 (Teorema da Convergência Monótona)**.** *Se (f_n) é uma sequência de funções de X em* R *mensuráveis e não negativas, a sequência é monotonamente crescente e converge a f, então*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Demonstração. (BARTLE, 1995, pág. 31) □

Teorema A.1.2 (Lema de Fatou)**.** *Se (f_n) for uma sequência de funções não negativas e Borel-mensuráveis de X em* R*, então*

$$\int_X (\liminf f_n) d\mu \leqslant \liminf \int_X f_n d\mu.$$

Demonstração. (BARTLE, 1995, pág. 33). □

Teorema A.1.3 (Teorema da Convergência Dominada)**.** *Se (f_n) for uma sequência de funções de X em* R *mensuráveis e integráveis, convergindo qtp para uma função f, e se existir g mensurável e integrável tal que $|f_n| \leqslant g$ $\forall n$, então f é integrável e*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Demonstração. (BARTLE, 1995, pág. 44) □

Teorema A.1.4. *Suponha que $\mu(X) < \infty$. Seja $f : X \to$ R *mensurável. Então*

$$||f||_1 \leqslant (\mu(X))^{1/2}||f||_2.$$

Demonstração. Pela desigualdade de Hölder,

$$\int_X |f| d\mu = \int_X 1|f| d\mu \leqslant (\int_X 1^2 d\mu)^{1/2} (\int_X |f|^2 d\mu)^{1/2} = (\mu(X))^{1/2}||f||_2.$$

□

A.2 Representação Canônica de Processos

Considere um espaço de probabilidade $(\Omega, \mathcal{F}, \mathrm{P})$ e seja $X : [0,T] \times \Omega \to \mathrm{R}^d$ um processo contínuo definido neste espaço. Considere o conjunto das funções contínuas $C([0,T]; \mathrm{R}^d)$, munido da sigma álgebra $\mathcal{C}$ gerada pelos cilindros

$$C = \{f : [0,T] \to \mathrm{R}^d \,|\, (f(t_1), \ldots, f(t_n)) \in B_1 \times \ldots B_n\}, \tag{A.1}$$

$0 \leqslant t_1 \leqslant \ldots \leqslant t_n \leqslant T, B_i$ boreliano em R. Mostra-se que a aplicação

$$\tilde{X} : \Omega \to C([0,T]; \mathrm{R}^d); \omega \to X_\omega : [0,T] \to \mathrm{R}^d; X_\omega(t) = X(t, \omega)$$

é mensurável, de forma que obtemos uma medida de probalidade P_X em $(C([0,T]; \mathrm{R}^d), \mathcal{C})$ que satisfaz

$$\mathrm{P}_X(C) = \mathrm{P}(X \in C), \qquad \forall C \in \mathcal{C}.$$

Esta medida de probabilidade é chamada de **representação canônica do processo** X, pois os processos X e $Y : [0,T] \times (C([0,T]; \mathrm{R}^d), \mathcal{C}, \mathrm{P}_X)$, onde

$$Y : [0,T] \times (C([0,T]; \mathrm{R}^d), \mathcal{C}, \mathrm{P}_X) \to \mathrm{R}^d; (t, \tilde{\omega}) \mapsto \tilde{\omega}(t),$$

são indistinguíveis, isto é, possuem as mesmas distribuições conjuntas. Por uma outra ótica, o teorema de extensão de Kolmogorov nos diz que a medida P_X induzida em $(C([0,T]; \mathrm{R}^d), \mathcal{C})$ é a única tal que (para um cilindro C do tipo (A.1))

$$\mathrm{P}_X(f \in C) = \mathrm{P}(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n).$$

E é bem por isso que X e $\tilde{X}$ são indistinguíveis, já que

$$\mathrm{P}_X(\tilde{X}_t \in B) = \mathrm{P}(X \in B). \tag{A.2}$$

Os detalhes são encontrados em (MARTIN; MARQUES, 1991). Havendo feito tal observação, será comum considerarmos processos como sua representação canônica. Dessa maneira, por exemplo, quando tivermos uma família de processos $X_{t,x}$, onde t, x são parâmetros da família (o caso mais comum será considerarmos $X_{t,x}$ como a única solução de uma equação diferencial estocástica, iniciada no ponto x no instante t, ou um movimento Browniano $W_{t,x}$, também iniciado em x no instante t), usaremos a notação $\mathrm{P}_{x,t}$ e $\mathrm{E}_{t,x}$, que deverão ser entendidas da seguinte forma: tomamos a representação canônica do processo X que se inicia em x no momento t e consideramos, em $(C([0,T]; \mathrm{R}^d), \mathcal{C}$ a medida de probabilidade induzida, $\mathrm{P}_{t,x}$. A partir daí, consideramos nossos processos ou variáveis aleatórias também definidas no espaço $(C([0,T]; \mathrm{R}^d), \mathcal{C}, \mathrm{P}_{t,x})$, para os quais podemos calcular, por exemplo, a respectiva esperança $\mathrm{E}_{t,x}$ em relação a essa medida de probabilidade.

A.3 Martingales

Teorema A.3.1 (do tempo de parada de Doob.)**.** *Seja* $\{M_t\}_{t=0}^T$ *um martingale contínuo com respeito à filtração* $\mathcal{F}_t$*, que satisfaze às condições usuais. Se* τ *for um tempo de parada para* $\mathcal{F}_t$*, então o processo*

$$X_t := M_{t\wedge\tau}$$

também é um martingale em relação a $\{\mathcal{F}_t\}$*.*

Demonstração. Enunciado e demonstração, ver (STEELE, 2001). □

O seguinte teorema é uma generalização deste primeiro.

Teorema A.3.2. *Seja* $\sigma \leqslant \tau$ *tempos de parada finitos. Seja* X *um martingale, submartingales ou supermartingale càd-làg. Então as v.a.'s* X_σ *e* X_τ *são integráveis e vale:*

- $X_\sigma = \mathrm{E}[X_\tau \mid \mathcal{F}_\sigma]$ *se* X *é martingale;*
- $X_\sigma \leqslant \mathrm{E}[X_\tau \mid \mathcal{F}_\sigma]$ *se* X *é submartingale;*
- $X_\sigma \geqslant \mathrm{E}[X_\tau \mid \mathcal{F}_\sigma]$ *se* X *é supermartingale;*

Demonstração. Ver (LOWTHER, 2009). □

Teorema A.3.3 (Desigualdades de Doob)**.** *Seja* M_t *um martingale ou um submartingale não negativo cujos caminhos são contínuos à direita com limites à esquerda. Então,*

1. $\mathrm{P}(\sup_{s\leqslant t} |M_s| \geqslant \lambda) \leqslant \mathrm{E}[|M_t|]/\lambda.$
2. *Se* $1 < p < \infty$*, então*

$$\mathrm{E}[\sup_{s\leqslant t} |M_s|^p] \leqslant \left(\frac{p}{p-1}\right)^p \mathrm{E}[|M_t|^p].$$

Demonstração. Ver (STEELE, 2001, Teorema 4.2) □

Definição A.3.4. *Dizemos que um processo* $X_t, t \in [0,T]$ *é uniformemente integrável se*

$$E[|\sup_{t\in[0,T]} X_t|] < \infty.$$

Teorema A.3.5. *Seja* M_t *um martingale local uniformemente integrável. Então* M_t *é um martingale.*

Demonstração. Seja τ_n uma sequência de tempos de parada tal que $\tau_n \to \infty$ qtp e $M_t^{\tau_n}$ é um martingale para todo n. Temos, para cada n,

$$M_s^{\tau_n} = \mathrm{E}[M_t^{\tau_n}|\mathcal{F}_s].$$

Tomando o limite quando $n \to \infty$, $M_s^{\tau_n} \to M_s$ e $M_t^{\tau_n} \to M_t$ qtp. Note que

$$\mathrm{E}[M_t^{\tau_n}] \leqslant \mathrm{E}[|\sup_{u\in[0,T]} M_u|] < \infty,$$

por hipótese, de forma que podemos aplicar o teorema da convergência dominada na primeira igualda para obter

$$M_s = \mathrm{E}[M_t|\mathcal{F}_s],$$

e o reusltado segue. □

Teorema A.3.6. *Seja M_t um martingale quadrado integrável em $[0,T]$. Então, para $0 \leqslant s < t \leqslant T$, temos*

$$\mathrm{E}[M_t^2 - M_s^2 \mid \mathcal{F}_s] = \mathrm{E}[(M_t - M_s)^2 \mid \mathcal{F}_s]$$

Teorema A.3.7 (Burkholder-Davis-Gundy)**.** *Para qualquer $0 < p < p$, existem constantes positivas c_p e C_p tais que, para qualquer martingale local contínuo M_t com $X_0 = 0$ e tempo de parada τ,*

$$c_p\mathrm{E}[\langle X\rangle_\tau^{p/2}] \leqslant \mathrm{E}[(X_\tau^*)^p] \leqslant C_p\mathrm{E}[\langle X\rangle_\tau^{p/2}]$$

Demonstração. Ver (LOWTHER, 2010). □

Teorema A.3.8 (Representação de Martingales)**.** *Seja $F \in L^2(\Omega \times [0,T])$ um processo adaptado à filtração Browniana n-dimensional. Então existe um único processo progressivamente mensurável $f \in L^2(\Omega \times [0,T]; \mathrm{R}^n)$ tal que*

$$F(\omega) = \mathrm{E}[F] + \int_0^T f(t,\omega)dW_t.$$

Demonstração. Ver (ØKSENDAL, 2010, Teorema 4.3.4) □