

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ADMINISTRAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO

Precificação e Hedge Dinâmico de Opções de Telebrás utilizando Redes Neurais

Moacir Arrieche Fernandes

Porto Alegre, 2000.

AGRADECIMENTOS

Agradeço ao professor Gilberto Kloeckner pela atenção dedicada a minha trajetória acadêmica, pela acolhida aos meus passos iniciais em finanças e aos conselhos que não se restringiram ao ambiente acadêmico mas que foram fundamentais para a construção deste trabalho.

RESUMO

As redes neurais podem ser uma alternativa aos modelos paramétricos tradicionais para a precificação de opções quando a dinâmica do ativo primário não for conhecida ou quando a equação associada à condição de não-arbitragem não puder ser resolvida analiticamente. Este trabalho compara a performance do modelo tradicional de Black-Scholes e as redes neurais. Os modelos foram utilizados para precificar e realizar a cobertura dinâmica das opções de compra das ações de Telebrás. Os resultados obtidos sugerem que as redes neurais deveriam ser consideradas pelos operadores de opções como uma alternativa aos modelos tradicionais.

ABSTRACT

Neural networks pricing formulas can be an alternative for the traditional parametric models when the underlying asset's price dynamics are unknown or when the no-arbitrage equation cannot be solved analytically. This study compares the performance of a neural network pricing formula with the traditional Black-Scholes. Both models were used to price and delta-hedge call options of Telebras. The results suggest that option traders should consider neural networks as an alternative for the traditional approach.

Sumário

AGRADECIMENTOS	3
RESUMO	4
ABSTRACT	5
1. INTRODUÇÃO	8
2. PROBLEMA, OBJETIVO E HIPÓTESES	11
2.1. Problema	11
2.2. Teste de Hipótese	13
2.3. Objetivos Gerais	13
2.4. Objetivos Específicos	14
3. JUSTIFICATIVA	15
4. MÉTODOS DE PRECIFICAÇÃO DE OPÇÕES	17
4.1. Abordagens Tradicionais	18
4.1.1. Modelos Estocásticos	18
4.1.1.1. Blocos Básicos (Wiener e Poisson)	18
4.1.1.2. Principais Modelos Estocásticos	22
4.1.1.3. Lema de Itô	26
4.1.2. Precificando Opções Através da Formação de Portfólios Livres de Risco	28
4.1.2.1. Construindo o portfolio	28
4.1.2.2. Aplicação – Modelo de Black-Scholes	30
4.1.3. Precificando Opções Utilizando Probabilidades Ajustadas	32
4.1.3.1. Obtenção da fórmula de Black-Scholes utilizando Martingales	34
4.1.4. Modelos Alternativos a Black-Scholes para Precificar Opções	39
4.1.4.1. Modelo de Merton (“Stochastic Interest Rate Model”)	39
4.1.4.2. Modelo Binomial	40
4.1.4.3. Modelo do Salto (“Pure Jump Model”)	42
4.1.4.4. Modelo do Salto-Difusão	42
4.1.4.5. Modelo de Volatilidade Estocástica	43
4.2. Modelos Não-Paramétricos	45
4.2.1. Valorização Canônica	46
4.2.2. Precificação de Opções Utilizando Redes Neurais	47
4.2.2.1. Redes Neurais	48
4.3. Evidências Empíricas	55
5. MÉTODO DE ESTUDO	59
5.1. A Rede	59
5.2. Fatores Utilizados	60
5.3. Amostra	63

5.4. Hedge Dinâmico	64
6. RESULTADOS	66
7. CONCLUSÕES	74
APÊNDICE 1: DETERMINAÇÃO DA VOLATILIDADE IMPLÍCITA	77
APÊNDICE 2: HEDGE DINÂMICO	79
REFERÊNCIAS E BIBLIOGRAFIA	82

1. Introdução

A teoria de precificação de opções tem seu principal marco na publicação do modelo de Black e Scholes (1973). A partir de algumas premissas como, por exemplo, o preço de um ativo segue um movimento browniano geométrico com volatilidade constante e ausência de custos de transação, Black e Scholes (B-S) conseguem mostrar que, se a volatilidade futura de determinado ativo é conhecida, o pagamento final de uma opção pode ser reproduzido utilizando-se uma estratégia de hedge dinâmico delta-neutro. Portanto, se não houver arbitragem, o preço da opção deve ser igual ao custo de reproduzir seu pagamento final.

A estratégia de hedge dinâmico delta-neutro, utilizada por B-S, consiste na imunização dos riscos decorrentes de pequenas variações no preço do ativo primário através da compra ou venda deste ativo. Desta forma, num portfolio delta-neutro, as variações no preço de uma opção devido a pequenas oscilações no preço do ativo primário são compensadas pela posição tomada neste ativo.

No entanto, o modelo possui algumas limitações e desvios. Os desvios mais conhecidos são o “*smile*” e a estrutura a termo da volatilidade. Com base nos preços negociados, pode-se verificar que a volatilidade implícita à fórmula de B-S varia conforme o preço de exercício (*smile*) e o prazo de vencimento (estrutura a termo da volatilidade). Hull (1997) afirma que, para contornar estes desvios, operadores utilizam matrizes de volatilidade que mostram a relação da volatilidade implícita com o preço de exercício e prazo de vencimento das opções.

Apesar dos desvios e limitações, B-S segue sendo o modelo utilizado pela maioria dos operadores de mercado, como afirma Britten-Jones (1998). Uma das explicações da preferência pela fórmula de B-S é dada por Taleb (1997):

“ É melhor aprimorar um modelo simples mas bem testado que operar com um modelo novo mais avançado.

É melhor usar um modelo com o menor número de parâmetros para estimar .” (pág. 109)

Cox e Rubinstein (1985) e Hull (1997) descrevem vários modelos que surgiram para sobrepor as limitações e desvios da fórmula de Black e Scholes. Modelos em que a volatilidade futura do ativo primário é incerta, modelos que permitem que o preço da ação dê “saltos” em vez de ter uma mudança contínua ou, ainda, modelos que consideram a ação como uma opção sobre os ativos de uma empresa.

Uma alternativa aos modelos paramétricos são os modelos que não restringem o movimento da ação a nenhum modelo estrutural. Por outro lado, como lembram Campbell, Lo e MacKinlay (1997), estes modelos não-paramétricos¹ apresentam uma desvantagem ao demandar uma substancial quantidade de dados para se ajustarem.

Hutchinson, Lo e Poggio (1994) utilizam uma destas ferramentas não-paramétricas, as redes neurais, para determinar o preço das opções de compra. Redes neurais são modelos matemáticos representados por uma coleção de simples unidades computacionais interligadas por um sistema de conexões. Este grande número de conexões e interligações entre seus elementos permite à rede um alto grau de adaptabilidade em aproximar funções.

¹ Apesar de serem chamados de modelos não-paramétricos, estes modelos contém parâmetros. Este trabalho apenas segue a classificação encontrada na literatura, modelos não-paramétricos, para os modelos em que a dinâmica do ativo não é determinada por uma equação diferencial estocástica

Devido a sua alta flexibilidade em aproximar relações complexas não-lineares, as redes neurais podem ser utilizadas para precificar derivativos que possuem fórmulas de precificação altamente não-lineares, mesmo quando estas estão disponíveis em soluções fechadas.

Se os fatores determinantes de determinada função estiverem disponíveis, pode-se treinar a rede para avaliar as relações entre as variáveis de entrada e saída, mesmo sem se entender completamente a dinâmica do processo. Tudo que se tem a fazer é definir o problema de forma a saber quais são os fatores que influenciam determinada variável e deixar que o sistema infira a relação entre as variáveis. As redes neurais são descritas com mais detalhes na seção 4.2.2.1.

Este estudo visa comparar a performance de uma rede neural e o modelo de B-S. Pretende-se utilizar a rede neural para a tomada de decisão de compra ou venda das opções de compra da Telebrás e comparar sua performance numa estratégia de hedge dinâmico delta-neutro com o modelo de B-S.

Esse trabalho foi dividido em 7 capítulos. O próximo capítulo contém os objetivos desse estudo; o terceiro apresenta argumentos para a realização deste estudo; o quarto analisa os métodos de precificação de opções; o quinto descreve o método de trabalho que foi utilizado; o sexto apresenta os resultados e o sétimo apresenta as conclusões do estudo.

2. Problema, Objetivo e Hipóteses

Este capítulo apresenta o problema a ser resolvido, os objetivos a serem alcançados e as hipóteses que serão testadas a fim de se responder o problema.

2.1. Problema

Para se tomar uma decisão de compra ou venda de uma opção, faz-se uma análise de *hedging*. Esta análise é, geralmente, baseada na comparação entre a volatilidade esperada do ativo e a volatilidade implícita nos preços de mercado. Rubin (1998) faz uma análise da arbitragem de volatilidade e propõe uma fórmula para o ganho esperado, conforme a visão particular do *trader* sobre o valor real da volatilidade que ocorrerá durante o período em que a posição for mantida.

Se o preço da opção for considerado muito baixo, compra-se a opção, eliminam-se os riscos e obtém-se um retorno livre de risco. Caso contrário, vende-se a opção e tenta-se reproduzir dinamicamente seu pagamento final a um custo inferior ao preço de mercado da opção.

Na maioria dos casos, a ferramenta utilizada para a tomada de decisão é o modelo de Black-Scholes. Este modelo não só é utilizado para precificação das opções, mas também orienta o processo de *hedge* dinâmico das opções.

Por outro lado, é amplamente reconhecido que este modelo é imperfeito em diversos aspectos. Assume que o ativo básico tem uma distribuição lognormal com uma volatilidade conhecida, ignora os custos de transação e assume que as negociações se dão de forma contínua.

Taleb (1997) mostra que, mesmo se os parâmetros que regulam o comportamento do ativo fossem conhecidos e que este ativo se comportasse da forma prevista pelo modelo, haveria uma flutuação dos resultados da cobertura dinâmica dos riscos da opção. Esta flutuação não poderia ser evitada uma vez que o *hedge* não pode ser feito de maneira contínua. Para diminuir a dispersão dos resultados, o *hedge* teria de ser feito o mais freqüentemente possível.

Na prática, esta eficiência na diminuição da dispersão dos resultados tem que ser contrabalançada com os custos de transação. O principal custo a ser analisado é o “*bid-ask spread*”, diferença entre o preço de oferta e procura do ativo. Cada vez que rebalancear o portfólio, metade da diferença entre o preço de compra e venda do ativo afetará negativamente o resultado final. Kamal e Derman (1999) fazem uma análise da estratégia de *hedge* dinâmico das opções e mostram que o desvio dos resultados é proporcional ao *vega* (derivada da opção em relação à volatilidade) da opção, se o ativo base se comportasse conforme as premissas de B-S.

Consequentemente, os agentes que negociam opções baseados em uma análise de volatilidade acabam enfrentando um risco adicional: ter que gerenciar sua exposição baseados em um modelo que não reflete as condições encontradas no mercado.

Assim, a formulação do problema estudado assumiu a seguinte fórmula:

O operador que utilizar uma rede neural, modelo construído a partir das condições encontradas no mercado, consegue executar uma cobertura de risco de suas posições de uma forma mais eficiente do que se utilizasse o modelo de Black e Scholes ?

2.2. Teste de Hipótese

Para se analisar a performance do modelo de B-S e a da rede neural, serão analisados os resultados de hedge dinâmico dos dois modelos. O resultado de hedge dinâmico é a diferença entre a venda (compra) da opção e o custo (resultado) de se replicar a opção dinamicamente.

As hipóteses a serem testadas para responder o problema são as seguintes:

- *Hipótese nula* (H0): Os resultados de hedge dinâmico do modelo de B-S e da rede neural são iguais
- *Hipótese Alternativa* (H1): Os resultados de hedge dinâmico da rede neural são superiores ao do modelo de Black e Scholes.

2.3. Objetivos Gerais

Esta dissertação visa contribuir para o entendimento dos fatores determinantes da formação de preço e, conseqüentemente, dos riscos associados às opções. Pretende-se descobrir a relação entre as variáveis de entrada da fórmula de Black e Scholes (taxa de juros, preço da ação, volatilidade do ativo, preço e data de exercício) através da construção de um modelo não-paramétrico: as redes neurais.

Adicionalmente, espera-se que este estudo divulgue esta ferramenta de análise de processos não lineares e incentivar pesquisadores e profissionais do mercado financeiro a testar a metodologia proposta em outras opções negociadas no mercado brasileiro, como opções sobre a taxa de câmbio e taxa de juros.

2.4. Objetivos Específicos

- a) Construir uma rede neural para precificar as opções de compra de Telebrás. A rede será treinada utilizando-se dados de mercado para que o modelo deduza as relações entre os fatores determinantes do preço destas opções.

- b) Utilizar a rede para a tomada de decisão de comprar ou vender as opções e verificar a performance do modelo construído através de uma estratégia de *hedge* dinâmico *delta-neutral*. Para avaliar a performance da rede neural, os resultados deste *hedge* dinâmico serão comparados com os que seriam obtidos utilizando-se o modelo de Black-Scholes.

3. Justificativa

Os instrumentos derivativos têm se tornado cada vez mais difundidos entre os agentes econômicos que o utilizam para controlar sua exposição a diversos fatores. Através do uso destes instrumentos, pode-se formar diversas estruturas que atendam às necessidades de empresas e investidores.

A compreensão do funcionamento destes instrumentos, assim como das variáveis que afetam seu valor é de fundamental importância para o processo de gerenciamento do risco financeiro. Quanto maior for o conhecimento sobre os instrumentos de *hedge*, e a sua sensibilidade às variáveis de mercado, mais eficiente será o controle do risco a que os investidores e empresas estão expostos.

O preço de uma opção de compra e a variação deste preço a seus fatores determinantes, normalmente, são calculados utilizando-se o modelo de Black-Scholes ou algum outro modelo que especifica o movimento do ativo primário da opção. Estes modelos são deduzidos a partir de premissas que nem sempre refletem a realidade dos mercados. Wilmott (1998) mostra como minimizar os efeitos do afastamento entre as condições ideais dos modelos e as condições de mercado através do ajuste dos parâmetros implícitos nos preços de mercado. Um dos ajustes sugeridos, provavelmente um dos mais utilizados, se dá através da construção de superfícies de volatilidade.

Uma outra forma de se construir um modelo se dá através da tentativa de se inferir as relações entre as variáveis através da observação de dados de mercado sem limitar o movimento do ativo básico a nenhuma equação. Desta forma, o modelo é construído a partir das condições prevalecentes no mercado ao invés de se partir de premissas e tentar adaptar o modelo às condições existentes.

Este estudo apresenta um modelo de precificação construído a partir de dados históricos que mostram a relação entre os fatores reconhecidos como determinantes do preço de uma opção de compra. Utilizou-se uma rede neural para inferir a relação entre as variáveis devido a sua utilidade quando não se tem certeza sobre a natureza de determinado processo, ou quando a relação entre as variáveis determinantes deste processo podem mudar com o passar do tempo, como descreve Haykin (1999). Outras características das redes neurais são discutidas na seção 4.2.2.1.

Os operadores de opções que utilizarem esta ferramenta poderão criar um modelo sem a necessidade de se determinar previamente como se dá a interação entre as variáveis de entrada e que pode se adaptar a mudanças nas relações entre estas variáveis que possam ocorrer ao longo do tempo. Portanto, podem gerenciar sua exposição baseados em um modelo que é determinado a partir das condições encontradas no mercado.

4. Métodos de Precificação de Opções

Como afirmam Campbell, Lo e MacKinlay (1997), a precificação de opções, bônus de subscrição (warrants) e outros derivativos—ativos financeiros cujo pagamento depende do preço de outros ativos— é um dos grandes sucessos da teoria financeira. Black e Scholes (1973), utilizando a condição de não-arbitragem, desenvolveram um modelo que teve uma aceitação quase imediata entre acadêmicos e profissionais do mercado financeiro.

A condição de não-arbitragem pode ser definida como ausência de oportunidades de arbitragem. Hull (1997) caracteriza uma oportunidade de arbitragem como a tomada de posição em diferentes ativos de tal forma que se possa garantir um rendimento maior que a taxa livre de risco. Então, oportunidades de arbitragem podem aparecer se for possível fazer uma série de investimentos sem desembolso líquido e, ainda, ter um lucro positivo esperado.

Campbell, Lo e MacKinlay (1997) classificam o método utilizado por Black e Scholes (1973) como paramétrico, pois utiliza uma equação diferencial estocástica, cujos parâmetros precisam ser estimados, para representar a dinâmica do ativo. Afirma também que foram as dificuldades em estimar modelos paramétricos da dinâmica dos preços dos ativos que levaram a várias tentativas de capturar esta dinâmica de uma forma não paramétrica. As redes neurais, objeto deste estudo, são uma forma não paramétrica.

Neste capítulo são analisados a abordagem tradicional (métodos paramétricos) de precificação de opções, a abordagem não paramétrica, com uma atenção especial para as redes neurais, e algumas evidências empíricas sobre a utilização destes modelos.

4.1. Abordagens Tradicionais

Neftci (1996) mostra que há dois modos tradicionais de se precificar derivativos. O primeiro, utilizando a noção de arbitragem, calcula probabilidades sob as quais os ativos se comportam como *martingales*² e, utilizando essas probabilidades ajustadas, determina-se o preço livre de arbitragem através da esperança implícita. O segundo método é mais direto. Consiste na construção de um portfólio livre de risco. Com a formação deste portfólio, chega-se a uma equação diferencial parcial implícita pela ausência de oportunidades de arbitragem. Esta equação pode ser resolvida de forma analítica ou numérica, resultando numa função que determina o preço do derivativo.

Os dois métodos utilizam modelos estocásticos, que são discutidos na próxima seção, para representar a dinâmica do ativo primário

4.1.1. Modelos Estocásticos³

Os modelos paramétricos geralmente se utilizam de equações diferenciais estocásticas para representar a dinâmica do ativo primário. O processo de Wiener e o de Poisson constituem os blocos básicos da modelagem de ativos financeiros em tempo contínuo. Combinando estes dois blocos básicos de forma apropriada, pode-se gerar modelos adequados a diversas aplicações.

4.1.1.1. Blocos Básicos (Wiener e Poisson)

O processo de Wiener, ou movimento browniano, é um processo estocástico contínuo que pode ser usado se os mercados forem dominados por eventos ordinários. O

² Um processo estocástico se comporta como um *martingale* quando sua trajetória não mostra nenhuma tendência ou periodicidade perceptível. (Neftci 1996)

³ O material apresentado nesta seção está baseado em Neftci (1996), Wilmott (1998), Hull (1997) e Cox e Rubinstein (1985). Uma maior profundidade do assunto pode ser encontrada em Oksendal (1992).

processo de Poisson pode ser usado para modelar saltos sistemáticos causados por eventos raros.

Processo de Wiener

Um processo de Wiener é adequado se a variável estocástica só pode mudar continuamente. O movimento browniano, ou processo de Wiener, é um tipo especial de processo estocástico de Markov. Um processo de Markov é um processo estocástico em que apenas o valor presente de uma variável é relevante para prever o futuro. Esta propriedade é consistente com a forma fraca de eficiência de mercado⁴, uma vez que o histórico da variável e o caminho percorrido para se chegar ao valor atual são irrelevantes.

Uma variável z , que segue um processo de Wiener, possui as seguintes propriedades:

1. Suas variações, Δz , estão relacionadas com um pequeno intervalo de tempo Δt através da equação

$$\Delta z = \epsilon \sqrt{\Delta t}$$

onde ϵ é uma variável aleatória de distribuição normal padronizada (média zero e desvio padrão unitário).

2. Os valores de Δz para dois intervalos diferentes de tempo Δt são independentes.

$$\text{Levando-se ao limite } \Delta t \rightarrow 0 : dz = \epsilon \sqrt{dt}$$

Da propriedade 1, vemos que Δz tem uma distribuição normal com média zero e variância Δt . A Figura 1 mostra um processo de Wiener desenvolvido em Excel.

⁴ A forma fraca de eficiência de mercado especifica que o preço de uma ação contém toda a informação baseada em preços passados

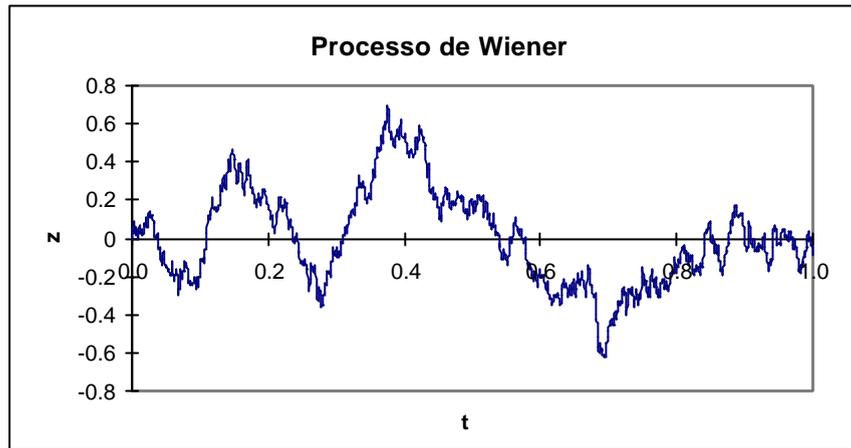


Figura 1 : Movimento Browniano (variância $\Delta t = 0.001$)

Processo de Poisson

Os preços dos ativos, às vezes, apresentam saltos ou grandes quedas inesperadamente. Para capturar este fenômeno, o processo de Poisson pode ser utilizado.

O processo de Poisson, z_t , pode ser descrito por uma variável cujos incrementos possuem apenas dois valores possíveis. Serão iguais a zero, significando que nenhum evento relevante ocorreu, ou será igual a um, caso algum evento significativo tenha ocorrido.

$$dz_t = \begin{cases} 1 & \text{com probabilidade } I dt \\ 0 & \text{com probabilidade } (1 - I) dt \end{cases}$$

A principal diferença do movimento browniano é que o tamanho do incremento não depende de dt . Ao invés disso, as probabilidades são funções de dt . A Figura 2 ilustra um processo de Poisson, enquanto a Figura 3 mostra uma mistura de Poisson com Wiener. Ambos processos foram gerados em Excel.

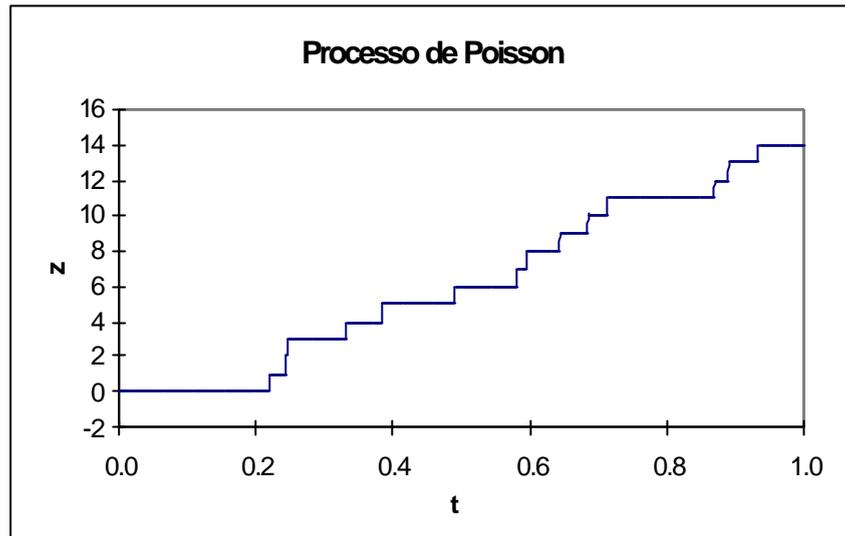


Figura 2: Processo de Poisson ($\Delta t = 0.001$; $I = 10$)

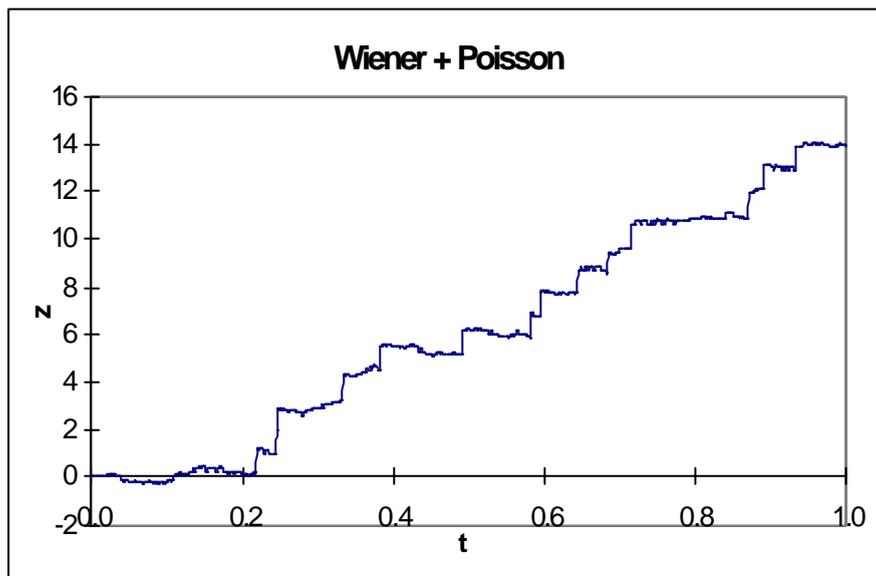


Figura 3: Poisson + Wiener ($\Delta t = 0.001$; $I = 10$)

Merton (1976) foi o primeiro a capturar a característica dos saltos nos preços dos ativos através do modelo salto-difusão. Wilmott (1998) cita três razões pelas quais, apesar de representar um fenômeno observável, modelos do tipo salto-difusão são raramente utilizados na prática. Estas razões seriam: dificuldade em estimar os parâmetros do modelo, solução e impossibilidade do hedge perfeito.

Partindo-se destes dois blocos básicos, *Wiener* e *Poisson*, pode-se gerar modelos adequados a diversas aplicações.

4.1.1.2. Principais Modelos Estocásticos

A maior parte dos modelos utilizados para representar a dinâmica dos preços dos ativos, S_t , possuem a seguinte forma

$$dS_t = a(S_t, t) dt + b(S_t, t) dz_t$$

onde dz_t é o termo de inovação representando eventos imprevisíveis que podem ocorrer em dt , conhecido como Processo de Wiener ou movimento browniano. O coeficiente $a(S_t, t)$ representa a tendência de S_t , ou a taxa de variação esperada por unidade de tempo, e $b(S_t, t)$, coeficiente de difusão, amplifica a incerteza do movimento de S_t .

Essa seção apresenta os principais modelos estocásticos utilizados na modelagem de ativos financeiros.

Movimento Browniano Aritmético

O modelo mais simples de equações diferenciais estocásticas ocorre quando os coeficientes de tendência e de difusão são constantes

$$dS_t = \mathbf{m} dt + \mathbf{s} dz_t$$

onde \mathbf{m} representa o coeficiente de tendência, \mathbf{s} o de difusão e z_t é um processo padrão de Wiener.

Para este processo, tem-se

$$E [dS_t] = \mathbf{m} dt$$

$$\text{Var} [dS_t] = \mathbf{s}^2 dt$$

Como pode ser observado na Figura 4, S_t parece flutuar ao redor de uma reta de inclinação m . Este modelo é adequado para ativos que apresentam tendência linear, cujas flutuações não dependam do nível de preços e não apresenta saltos.

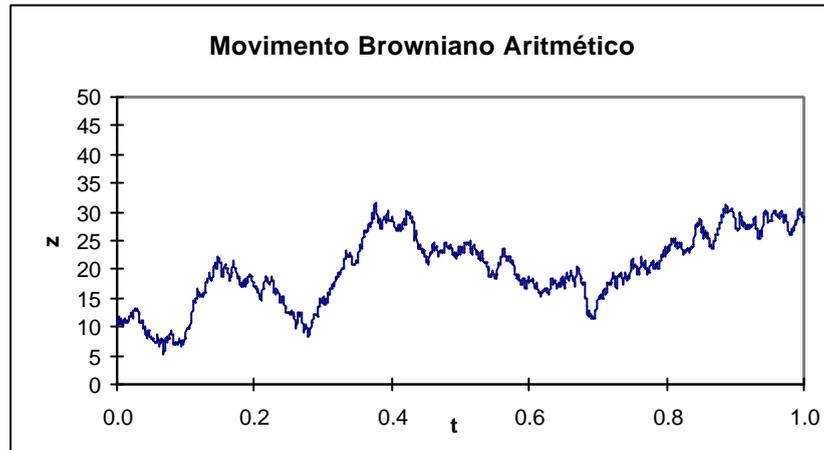


Figura 4: Movimento Browniano Aritmético ($m = s = 20$)

Outro ponto a ser observado é que o modelo permite que S_t assumam valores negativos. Conseqüentemente, o modelo aritmético não é muito adequado para representar a dinâmica de ativos como taxa de juros e preço de ações. Os modelos apresentados, a seguir, não permitem que S_t assumam valores negativos se seu valor inicial for positivo.

Movimento Browniano Geométrico e o Processo da Raiz Quadrada

O modelo mais aceito para modelar o comportamento de preço de ações, moedas, commodities e índices é o modelo geométrico (Wilmott, 1998). Este foi o modelo adotado por Black e Scholes (1973) onde a tendência e o desvio são proporcionais a S_t .

$$dS_t = mS_t dt + s S_t dz_t$$

Enquanto o modelo aritmético descreve o comportamento de um ativo que flutua ao redor de uma linha reta, este modelo nos dá preços que flutuam ao redor de uma tendência exponencial. A Figura 5 ilustra o processo.

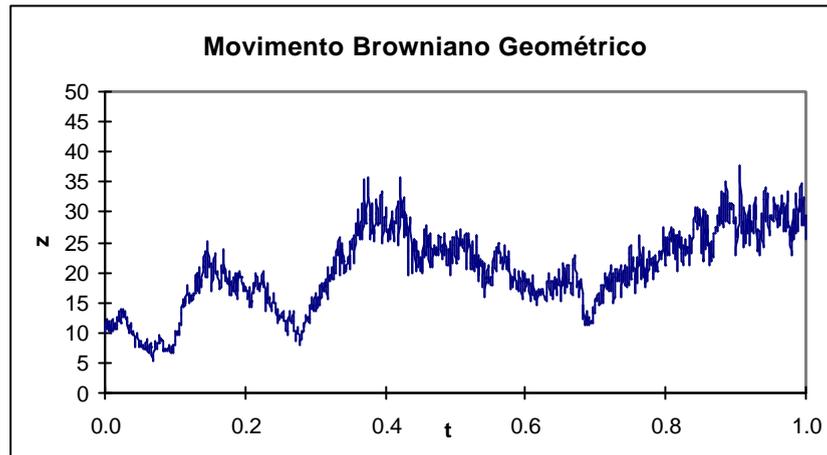


Figura 5: Movimento Browniano Geométrico ($m=20$ $s=3$)

Note que a variância de $S_t - S_{t-1}$ é proporcional ao quadrado do ativo.

$$\text{Var}(S_t - S_{t-1}) = s^2 S_{t-1}^2$$

Isto pode acrescentar muita variação ao preço do ativo. Um modelo similar, em que o ativo apresenta menos flutuações, é o modelo da raiz quadrada

$$dS_t = \mathbf{m} S_t dt + s \sqrt{S_t} dz_t$$

como podemos ver na Figura 6.

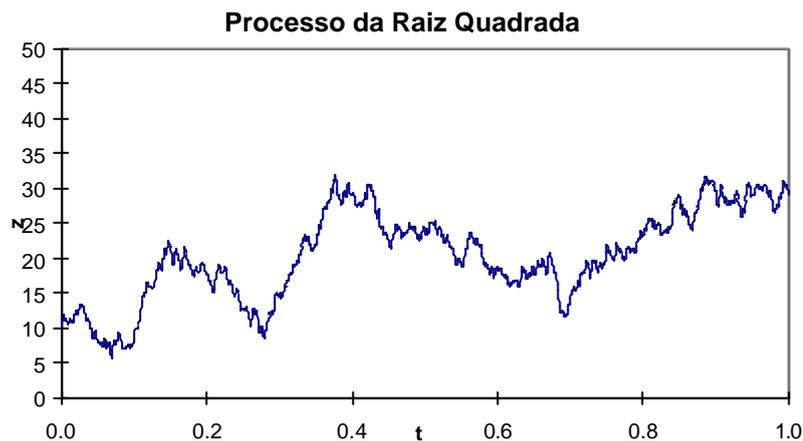


Figura 6: Movimento da Raiz Quadrada ($m=20$ $s=3$)

O próximo modelo é bastante utilizado para derivativos de taxas de juros.

Processo de Reversão à Média (Mean Reverting)

Esse modelo é representado pela seguinte equação diferencial estocástica

$$dS_t = \mathbf{I} (\mathbf{m} - S_t) dt + \mathbf{s} S_t dz_t$$

De acordo com este modelo, assim que o valor do ativo se distancia de um valor médio \mathbf{m} , ele tende a reverter para a média novamente. A velocidade de reversão é controlada pelo parâmetro $\mathbf{I} > 0$. Pela equação do modelo, se S_t estiver abaixo da média, o termo entre parênteses ($\mathbf{m} - S_t$) se torna positivo fazendo com que seja mais provável que dS_t seja positivo também. A Figura 7 ilustra o processo.

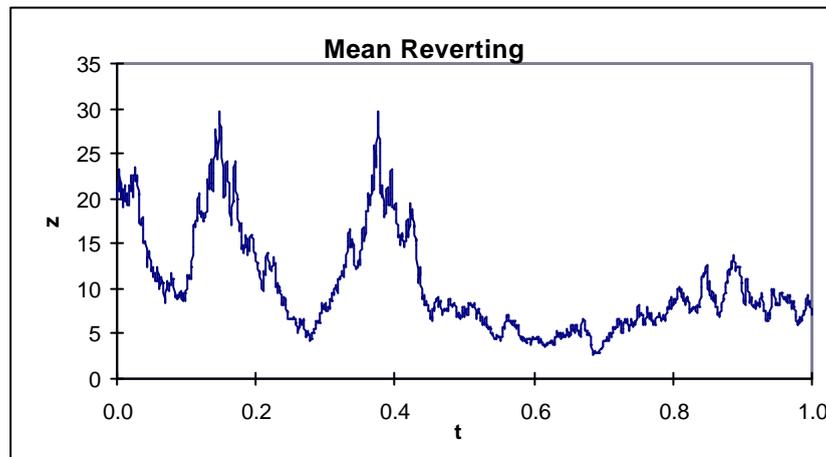


Figura 7: Processo de Reversão à Média (*Mean Reverting*: $\mathbf{m}=10$; $\mathbf{s}=2$; $\mathbf{I}=5$)

Hull (1997) apresenta alguns argumentos econômicos que justificam a preferência desse modelo na precificação de derivativos de taxa de juros. Quando as taxas de juros se encontram em patamares elevados, a economia tende a desacelerar resultando numa menor demanda por recursos. Conseqüentemente, as taxas caem. Quando as taxas de juros estão baixas, ocorre uma maior procura por recursos que resulta na elevação das taxas.

Volatilidade Estocástica

Nos modelos anteriores, o parâmetro de difusão é constante ou proporcional a S_t . Um modelo mais amplo pode ser obtido permitindo que o parâmetro de difusão também seja aleatório.

Um dos modelos de volatilidade estocástica mais utilizado, é o modelo em que a volatilidade segue um processo de reversão à média

$$dS_t = \mathbf{m}S_t dt + \mathbf{s}_t dz_1$$

$$d\mathbf{s}_t = \mathbf{I} (\mathbf{s}_o - \mathbf{s}_t) dt + \mathbf{u} \mathbf{s}_t dz_2$$

onde os processos de Wiener, z_1 e z_2 , podem ser correlacionados.

Neste modelo a volatilidade pode se desviar de sua média a longo prazo, \mathbf{s}_o , revertendo com uma velocidade ajustada pelo parâmetro $\mathbf{I} > 0$. O parâmetro $\mathbf{u} > 0$ indica a dispersão da volatilidade. A Figura 8 ilustra o processo.

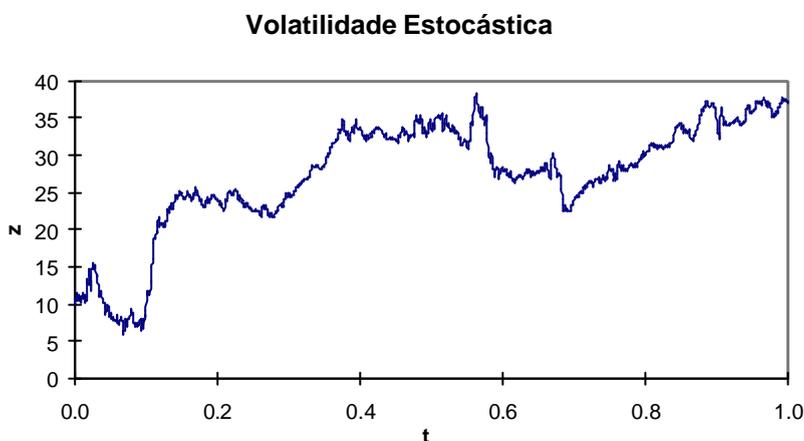


Figura 8: Modelo de Volatilidade Estocástica

Todos os modelos aqui apresentados possuem uma fonte de indeterminação, representada pela variável z_t , o processo de *Wiener*. Neftci (1996) mostra que a variável z_t tem um processo muito irregular, que não possui uma derivada determinística. Portanto, para se determinar incrementos em S_t , que depende de z_t , tem-se que utilizar o Lema de Itô.

4.1.1.3. Lema de Itô

Suponha que $f(x)$ seja uma função do processo aleatório x . Poderíamos representar $f(x)$ através de sua expansão de Taylor em torno de um valor conhecido de x , x_0 .

$$f(x) = f(x_0) + f_x(x_0)[x - x_0] + \frac{1}{2} f_{xx}(x_0)[x - x_0]^2 + \frac{1}{3!} f_{xxx}(x_0)[x - x_0]^3 + R(x, x_0) \text{ onde}$$

$R(x, x_0)$ representa os termos de ordens superiores da série de Taylor.

Se x fosse uma variável determinística, à medida que $\Delta x (= x - x_0)$ se aproximasse de zero poderíamos desprezar os termos superiores

$$f(x) = f(x_0) + f_x(x_0)\Delta x$$

No caso de um processo aleatório, o mesmo não pode ser feito. Por mais que se diminua a distância entre x e x_0 , não podemos desprezar seu termo de segunda ordem porque x tem variância positiva, isto é,

$$E[\Delta x]^2 > 0$$

Tratando Δx^2 como zero, seria o mesmo que afirmar que x tem variância nula. Neftci (1996) mostra que, para um processo aleatório, os incrementos de f podem ser determinados pelo Lemma de Itô:

Seja $F(S_t, t)$ uma função diferenciável de t e do processo aleatório S_t

$$dS_t = a_t dt + \mathbf{s}_t dz_t$$

Então, tem-se

$$dF_t = \frac{\partial F}{\partial S_t} dS_t + \frac{\partial F}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \mathbf{s}_t^2 dt$$

ou, após substituir dS_t

$$dF_t = \left[\frac{\partial F}{\partial S_t} a_t + \frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \mathbf{s}_t^2 \right] dt + \frac{\partial F}{\partial S_t} \mathbf{s}_t dz_t$$

onde a igualdade é válida no sentido de convergência quadrática, isto é, quando $dt \rightarrow 0$, as variâncias se igualam.

Portanto, utilizando-se o Lemma de Itô, pode-se relacionar variações no preço do derivativo com base nas variações do ativo primário. Como as duas variações são motivadas pela mesma fonte de incerteza, Hull (1997) mostra que é possível formar portfólios livres de risco em tempo contínuo. A próxima seção ilustra o processo.

4.1.2. Precificando Opções Através da Formação de Portfólios Livres de Risco

Uma das formas tradicionais de se determinar o preço de um derivativo se dá através da construção de um portfólio livre de risco composto pelo ativo primário e o próprio derivativo. A construção deste portfólio só é possível porque o derivativo e o ativo primário são afetados pela mesma fonte de incerteza. Black e Scholes (1973) foram os primeiros a utilizar este método para chegar ao preço das opções de compra.

Com a formação deste portfólio, chega-se a uma equação diferencial parcial implícita pela ausência de oportunidades de arbitragem. Esta equação pode ser resolvida de forma analítica ou numérica, resultando numa função que determina o preço do derivativo.

4.1.2.1. Construindo o portfolio

Seja $F(S_t, t)$ o preço do derivativo e S_t o preço do ativo básico, governado pela seguinte diferencial estocástica

$$dS_t = a_t dt + \mathbf{s}_t dz_t$$

Então, tem-se

$$dF_t = \frac{\partial F}{\partial S_t} dS_t + \frac{\partial F}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \mathbf{s}_t^2 dt$$

Se um portfólio P_t é composto por uma unidade do derivativo e $-\frac{\partial F}{\partial S}$ unidades

do ativo primário, tem-se

$$\begin{aligned} dP_t &= dF - \frac{\partial F}{\partial S_t} dS_t \Rightarrow \\ dP_t &= \frac{\partial F}{\partial S_t} dS_t + \frac{\partial F}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \sigma_t^2 dt - \frac{\partial F}{\partial S_t} dS_t \Rightarrow \\ dP_t &= \frac{\partial F}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \sigma_t^2 dt \end{aligned}$$

Como pode-se observar, o portfólio P_t não possui nenhum termo aleatório, sendo totalmente previsível. Isto significa que P_t é livre de risco. Portanto seu retorno deve ter uma taxa de retorno igual à taxa livre de risco, ou existirá uma oportunidade de arbitragem. Assumindo que a taxa livre de risco é dada por r , o retorno de P_t deve igualar

$$r P_t dt$$

Combinando as duas últimas expressões, pode-se escrever:

$$r P_t dt = \frac{\partial F}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \sigma_t^2 dt$$

Eliminando-se dt e substituindo-se P_t chega-se à equação diferencial:

$$-rF + r \frac{\partial F}{\partial S_t} S_t + \frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \sigma_t^2 = 0$$

que deve ser satisfeita pelo derivativo.

No caso da opção de compra europeia, com preço de exercício K , sabe-se com certeza a relação entre o preço da opção no vencimento e o preço do ativo:

$$F(t = T) = [\max(S_t - K, 0)]$$

Consequentemente, o valor da opção pode ser determinado através de sua equação diferencial parcial e condições de contorno associadas. A obtenção da fórmula de Black e Scholes ilustra bem o processo.

4.1.2.2. Aplicação – Modelo de Black-Scholes

Black e Scholes (1973) foram os primeiros a perceber que poderiam formar um portfolio livre de risco composto por opções e ações. O que lhes permitiu a criação deste portfolio é o fato do preço da ação e da opção serem afetados pela mesma fonte de incerteza: o movimento da ação.

Para chegar à equação diferencial que resulta no preço da opção, Black e Scholes partem de algumas premissas:

- o preço da ação segue um movimento browniano geométrico com σ constante;
- é permitida a venda a descoberto dos ativos;
- não há custos de transação ou impostos;
- todos os ativos são perfeitamente divisíveis;
- não há dividendos durante o prazo da opção;
- a taxa livre de risco é constante independentemente do prazo;
- e não existem condições de arbitragem.

Partindo da equação do movimento browniano geométrico, para o movimento de uma ação

$$dS_t = \mathbf{m}S_t dt + \mathbf{s} S_t dz_t$$

chegam à correspondente equação diferencial de uma opção de compra, c , aplicando-se o Lemma de Itô:

$$dc = \left(\frac{\partial c}{\partial S} \mathbf{m}S + \frac{\partial c}{\partial t} + \frac{1}{2} \frac{\partial^2 c}{\partial S^2} \mathbf{s}^2 S^2 \right) dt + \frac{\partial c}{\partial S} \mathbf{s} S dz .$$

Tomando posições contrárias na ação e na opção de compra, montam um portfólio P que elimina a incerteza do processo de Wiener. Eliminada a incerteza, o retorno deve ser igual a taxa livre de risco, r :

$$dP = r P dt \quad .$$

Este portfólio P tem a seguinte composição:

➤ venda de uma opção de compra (-c)

➤ compra de $+\frac{\partial c}{\partial S}$ ações

Portanto, a variação dP pode ser representada por :

$$d\Pi = \left(-\frac{\partial c}{\partial t} - \frac{1}{2} \frac{\partial^2 c}{\partial S^2} \sigma^2 S^2 \right) dt = r\Pi dt \quad .$$

Substituindo-se P e eliminando-se dt chegam à equação diferencial

$$\frac{\partial c}{\partial t} + rS \frac{\partial c}{\partial S} + \frac{1}{2} \frac{\partial^2 c}{\partial S^2} \sigma^2 S^2 = rc$$

No caso da opção de compra europeia, têm-se as seguintes condições de contorno:

$$c = \max(S - K, 0) \quad \text{para } t = T, \text{ data de vencimento}$$

$$c = 0 \quad \text{para } S = 0$$

$$c \sim S \quad \text{quando } S \gg K$$

A solução do problema, dado que σ e r são constantes, é dada por:

$$c = S N(d_1) - K e^{-r(T-t)} N(d_2)$$

onde

$$d_1 = \frac{\ln(S/K) + (r + s^2/2)(T-t)}{s\sqrt{T-t}}$$

$$d_2 = d_1 - s\sqrt{T-t}$$

onde $N(d)$ é a função de probabilidade acumulada da distribuição normal, ou seja, representa a probabilidade da variável aleatória x ser igual ou inferior a d :

$$N(d_i) = \int_{-\infty}^{d_i} \frac{1}{\sqrt{2p}} e^{-\frac{1}{2}x^2} dx$$

A próxima seção mostra a outra forma tradicional de se precificar derivativos através da utilização de probabilidades ajustadas.

4.1.3. Precificando Opções Utilizando Probabilidades Ajustadas⁵

A outra forma tradicional de se calcular o preço de derivativos é através da transformação do processo do ativo primário em *martingale*. Neftci (1996) mostra que ambos os métodos dão a mesma resposta. Dependendo do problema a ser resolvido, o investidor pode escolher o método mais conveniente a ser utilizado.

Suponha que, num dado instante t , toda informação disponível possa ser representada por I_t . Uma variável aleatória X_t que satisfaça à igualdade

$$E^P[X_{t+s} | I_t] = X_t \quad \text{para todo } s > 0$$

é chamado de um martingale com relação à probabilidade P .

Se, ao invés disso,

$$E^{P'}[X_{t+s} | I_t] \geq X_t \quad \text{para todo } s > 0$$

⁵ Esta seção segue a linha de trabalho de Neftci (1996).

tem-se um submartingale com relação à probabilidade P' .

Ativos de risco, quando trazidos a valor presente pela taxa livre de risco, podem ser considerados submartingales sob a probabilidade real, pois normalmente satisfazem à desigualdade

$$S(t) < \frac{1}{(1+r)} E[S(t+1)]$$

para compensar o investidor pelo risco assumido. Caso contrário, o investidor não teria motivos para possuir determinado ativo em seu portfólio.

Mas nada impede que se possa calcular probabilidades ajustadas de tal forma que

$$S(t) = \frac{1}{(1+r)} E^P[S(t+1)]$$

onde $E^P[S(t+1)]$ é o valor esperado da variável S em $t+1$ calculado com probabilidade P .

Estas equações são convenientes porque internalizam qualquer prêmio de risco. De fato, não é necessário calcular prêmios de risco se utilizarmos probabilidades ajustadas.

Transformando submartingales sob probabilidades reais em martingales utilizando probabilidades ajustadas pode-se precificar derivativos seguindo as seguintes etapas:

1. obter um modelo que represente a dinâmica do ativo primário;
2. avaliar como o preço do derivativo é determinado no vencimento ou em outras fronteiras;
3. obter probabilidades ajustadas ao risco;
4. calcular o valor esperado do derivativo utilizando estas probabilidades;

5. descontar esta esperança utilizando a taxa livre de risco.

No caso da opção de compra européia, encontra-se a probabilidade ajustada P' sob a qual $S(t)$ se torna um martingale. Então, calcula-se:

$$c = E^{P'} e^{-r(T-t)} [\max(S_t - K, 0)]$$

analiticamente ou numericamente.

A seção a seguir ilustra bem o processo.

4.1.3.1. Obtenção da fórmula de Black-Scholes utilizando Martingales

A fórmula de Black-Scholes também pode ser obtida utilizando-se o martingale equivalente. Neste caso, o valor da opção será o valor esperado do resultado final, calculado com probabilidades ajustadas, trazidos a valor presente pela taxa livre de risco.

$$c = E^{P'} e^{-r(T-t)} [\max(S_t - K, 0)]$$

Considere que S_t siga o processo geométrico de Wiener

$$Y_t \sim N(\mathbf{m}t, \mathbf{s}^2 t)$$

$$S_t = S_0 e^{Y_t}$$

onde Y_t é um processo de tempo contínuo que segue uma distribuição normal.

Portanto, o valor esperado de e^{Y_t} pode ser calculado explicitamente:

$$E[e^{Y_t}] = \int_{-\infty}^{+\infty} e^{Y_t} \frac{1}{\sqrt{2\pi s^2 t}} e^{-\frac{1(Y_t - \mathbf{m})^2}{2 s^2 t}} dY_t$$

$$E[e^{Y_t}] = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi s^2 t}} e^{-\frac{1(Y_t - \mathbf{m})^2}{2 s^2 t} + Y_t} dY_t$$

Pode-se transformar o expoente num quadrado perfeito multiplicando-se o termo dentro da integral por

$$e^{-\left(mt + \frac{1}{2}s^2 t\right)} e^{\left(mt + \frac{1}{2}s^2 t\right)} = 1$$

Assim, tem-se

$$E \left[e^{Y_t} \right] = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi s^2 t}} e^{\left(mt + \frac{1}{2}s^2 t\right)} e^{-\frac{1}{2} \frac{(Y_t - mt)^2}{s^2 t} + Y_t - \left(mt + \frac{1}{2}s^2 t\right)} dY_t$$

$$E \left[e^{Y_t} \right] = e^{\left(mt + \frac{1}{2}s^2 t\right)} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi s^2 t}} e^{-\frac{1}{2} \frac{(Y_t - (mt + s^2 t))^2}{s^2 t}} dY_t$$

A integral da equação acima é igual a 1, já que representa a área sob a Distribuição

Normal. Portanto,

$$E \left[e^{Y_t} \right] = e^{\left(mt + \frac{1}{2}s^2 t\right)}$$

Determinando P'(probabilidade ajustada)

Para calcular o preço da opção precisamos achar a probabilidade ajustada P' de tal forma que

$$E \left[e^{-rt} S_t \mid S_u, u < t \right] = e^{-ru} S_u$$

para que S_t seja um martingale.

Relembrando,

$$S_t = S_0 e^{Y_t}; \quad Y_t \sim N(mt, s^2 t)$$

$$Y_t = Y_s + \int_s^t dY_u$$

$$\Delta Y_t = \int_s^t dY_u$$

$$\Delta Y_t \sim N(\mathbf{m}(t-s), \mathbf{s}^2 (t-s))$$

$$E[e^{\Delta Y_t}] = e^{\left(m(t-s) + \frac{1}{2}s^2(t-s)\right)} \Rightarrow$$

$$E[S_t | S_u, u < t] = S_u e^{\left(m(t-s) + \frac{1}{2}s^2(t-s)\right)}$$

Seja P' dada por

$$N(\mathbf{r}t, \mathbf{s}^2 t)$$

onde a única diferença entre P e P' é o viés (*drift*).

O valor esperado, dada a distribuição P' é

$$E^{P'}[e^{-r(t-u)} S_t | S_u, u < t] = [S_u e^{-r(t-u)}] e^{\left(r(t-u) + \frac{1}{2}s^2(t-u)\right)}$$

Pode-se selecionar \mathbf{r} de modo que P' satisfaça a condição de um martingale. Neste caso,

$$\mathbf{r} = r - \frac{1}{2}\mathbf{s}^2$$

satisfaz a condição.

Utilizando o valor de \mathbf{r} para calcular o valor esperado, tem-se

$$-r(t-u) + \mathbf{r}(t-u) + \frac{1}{2}\mathbf{s}^2(t-u) = 0 \Rightarrow$$

$$E^{P'}[e^{-r(t-u)} S_t | S_u, u < t] = S_u$$

satisfazendo a condição de martingale.

Portanto, sob a distribuição

$$N\left(\left(r - \frac{1}{2}\mathbf{s}^2\right)t, \mathbf{s}^2 t\right)$$

os ativos financeiros têm a propriedade de um martingale.

Tendo distribuição ajustada, pode-se calcular o preço do derivativo.

$$C = E^{P'} e^{-r(T-t)} [\max(S_t - K, 0)]$$

Pode-se escrever também

$$c = \int_{-\infty}^{+\infty} e^{-rt} \max[S_T - K, 0] dP'$$

Substituindo-se a distribuição de probabilidade ajustada e S_t ,

$$dP' = \frac{1}{\sqrt{2ps^2T}} e^{-\frac{1}{2s^2T}\left(Y_T - \left(r - \frac{1}{2}s^2T\right)\right)^2} dY_T$$

$$c = \int_{-\infty}^{+\infty} e^{-rt} \max[S_0 e^{Y_T} - K, 0] \frac{1}{\sqrt{2ps^2T}} e^{-\frac{1}{2s^2T}\left(Y_T - \left(r - \frac{1}{2}s^2T\right)\right)^2} dY_T$$

Mudando os limites da integral, consegue-se eliminar a função de máximo. Note que, a condição

$S_0 e^{Y_T} \geq K$ é equivalente a $Y_T \geq \ln(K/S_0)$. Logo,

$$c = \int_{-\ln\left(\frac{K}{S_0}\right)}^{+\infty} e^{-rt} (S_0 e^{Y_T} - K) \frac{1}{\sqrt{2ps^2T}} e^{-\frac{1}{2s^2T}\left(Y_T - \left(r - \frac{1}{2}s^2T\right)\right)^2} dY_T$$

$$c = A - B$$

$$A = S_0 \int_{-\ln\left(\frac{K}{S_0}\right)}^{+\infty} e^{-rt} e^{Y_T} \frac{1}{\sqrt{2ps^2T}} e^{-\frac{1}{2s^2T}\left(Y_T - \left(r - \frac{1}{2}s^2T\right)\right)^2} dY_T$$

$$B = Ke^{-rT} \int_{-\ln\left(\frac{K}{S_0}\right)}^{+\infty} e^{-rt} \frac{1}{\sqrt{2ps^2T}} e^{-\frac{1}{2s^2T}\left(Y_T - \left(r - \frac{1}{2}s^2T\right)\right)^2} dY_T$$

Aplicando a transformação

$$Z = \frac{Y_T - \left(r - \frac{1}{2}s^2\right)T}{s\sqrt{T}}$$

tem-se

$$B = Ke^{-rT} \int_{\frac{\ln\left(\frac{K}{S_0}\right) - \left(r - \frac{1}{2}s^2\right)T}{s\sqrt{T}}}^{+\infty} \frac{1}{\sqrt{2p}} e^{-\frac{1}{2}Z^2} dZ$$

O limite inferior da integral é igual a $-d_2$ (o parâmetro de Black e Scholes).

Portanto,

$$B = Ke^{-rT} \int_{-d_2}^{+\infty} \frac{1}{\sqrt{2p}} e^{-\frac{1}{2}Z^2} dZ$$

Utilizando-se uma das propriedades simétricas da Normal, tem-se

$$\begin{aligned} \int_{-L}^{+\infty} f(x)dx &= \int_{-\infty}^{+L} f(x)dx \Rightarrow \\ B &= Ke^{-rT} \int_{-\infty}^{d_2} \frac{1}{\sqrt{2p}} e^{-\frac{1}{2}Z^2} dZ = Ke^{-rT} N(d_2) \end{aligned}$$

Utilizando a transformação Z novamente,

$$\begin{aligned} A &= S_0 \int_{-\ln\left(\frac{K}{S_0}\right)}^{+\infty} e^{-\pi} e^{Y_T} \frac{1}{\sqrt{2ps^2T}} e^{-\frac{1}{2s^2T}\left(Y_T - \left(r - \frac{1}{2}s^2\right)T\right)^2} dY_T \\ A &= e^{\left(r - \frac{1}{2}s^2\right)T} e^{-rT} S_0 \int_{-d_2}^{+\infty} e^{sz\sqrt{T}} \frac{1}{\sqrt{2p}} e^{-\frac{1}{2}Z^2} dZ \\ A &= e^{\left(r - \frac{1}{2}s^2\right)T} e^{-rT} S_0 \int_{-\infty}^{+d_2} \frac{1}{\sqrt{2p}} e^{-\frac{1}{2}(Z^2 + 2sZ\sqrt{T})} dZ \end{aligned}$$

Formando um quadrado perfeito adicionando e subtraindo $\frac{s^2T}{2}$, temos

$$A = e^{\left(r - \frac{1}{2}s^2\right)T} e^{-rT} S_0 e^{\frac{Ts^2}{2}} \int_{-\infty}^{+d_2} \frac{1}{\sqrt{2p}} e^{-\frac{1}{2}(Z+s\sqrt{T})^2} dZ$$

Simplificando-se a expressão e realizando a substituição

$$H = Z + s\sqrt{T}$$

obtem-se

$$A = S_0 \int_{-\infty}^{+d_2 + s\sqrt{T}} \frac{1}{\sqrt{2p}} e^{-\frac{1}{2}H^2} dH = S_0 N(d_1)$$

onde

$$d_1 = d_2 + s\sqrt{T}$$

Portanto, consegue-se chegar à mesma solução de Black-Scholes sem resolver uma equação diferencial parcial.

A próxima seção mostra alguns exemplos de modelos paramétricos alternativos ao modelo de Black e Scholes para determinar o preço das opções.

4.1.4. Modelos Alternativos a Black-Scholes para Precificar Opções⁶

Essa seção mostra alguns dos modelos paramétricos que tentam sobrepor algumas das limitações do modelo de Black e Scholes.

4.1.4.1. Modelo de Merton (“Stochastic Interest Rate Model”)

O modelo de Merton (1973) pode ser considerado como um aperfeiçoamento do modelo de Black-Scholes que permite uma variação estocástica da taxa de juros. Seja $P(t, T)$ o valor de um bônus em t que paga \$1 no tempo T . O modelo de Merton assume que $P(t, T)$ segue o seguinte processo estocástico:

$$\frac{dP(t, T)}{P(t, T)} = \mathbf{m}_p dt + \mathbf{s}_p dz_p$$

onde a variável \mathbf{m}_p é a taxa de crescimento do preço do bônus, que pode ser estocástica; \mathbf{s}_p representa a volatilidade de $P(t, T)$, e assume-se que seja uma função conhecida dependendo apenas de t e T ; dz_p é um processo de Wiener.

Incorporando este processo no modelo de Black-Scholes, Merton chega a uma solução para o preço de uma opção de compra européia. Substituindo a taxa de juros livre

de risco por $R(t)$, a taxa contínua de juros do bônus $P(t, T) = e^{-R(t, T)(T-t)}$, e \mathbf{s} por \mathbf{d} , Merton chega a uma solução similar à de Black e Scholes:

$$c = SN(d_1) - P(t, T) XN(d_2)$$

onde

$$d_1 = \frac{\ln(S / X) - \ln P(t, T) + \mathbf{d}^2 (T - t) / 2}{\mathbf{d} \sqrt{T - t}}$$

$$d_2 = d_1 - \mathbf{d} \sqrt{T - t}$$

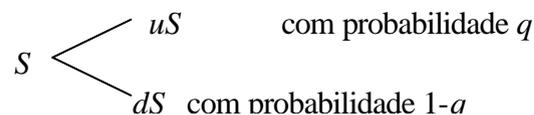
$$\mathbf{d}^2 (T - t) = \int_t^T (\mathbf{s}^2 + \mathbf{s}_p^2 - 2\mathbf{r}\mathbf{s}_p\mathbf{s}) dt$$

onde \mathbf{s} representa a volatilidade da ação e \mathbf{r} a correlação instantânea entre o preço do bônus e da ação. Hull (1997) afirma que, na prática, \mathbf{s}_p tende a ser bem menor que \mathbf{s} , aproximando desta forma \mathbf{d} de \mathbf{s} . Portanto, o ajuste de volatilidade tem um pequeno impacto no cálculo do preço da opção no Modelo de Merton.

4.1.4.2. Modelo Binomial

Cox, Ross e Rubinstein (1979) apresentam uma aproximação simplificada para a precificação de opções. Como na construção do modelo de Black e Scholes, obtêm o preço da opção através de uma condição de arbitragem.

O modelo binomial assume que o preço da ação, S , a cada pequeno intervalo de tempo, só pode ter dois valores: uS e dS , onde $u > 1$ e $0 < d < 1$.



A idéia atrás do modelo binomial é a formação de um portfolio cujo valor seja idêntico, independente da direção do movimento do preço da ação. Uma vez que o valor do portfolio, no final do período, é sabido com certeza, o retorno do mesmo só pode ser

⁶ Wilmott (1998), Hull (1997) e Cox e Rubinstein (1985) são exemplos de fontes de modelos paramétricos de opções. Smithson e Song (1998) apresentam um sumário das modificações e extensões do modelo de

igual à taxa livre de risco, r . Este portfolio pode ser construído através da compra de um número D de ações e venda de uma opção de compra.

Imagine o caso de uma opção de compra, c , onde:

S (preço da ação) = 100; K (preço de exercício) = 100; $d = 0,8$ e $u = 1,2$. Logo

$$100D - c \begin{cases} 120\Delta - c_u; (c_u = 10) \\ 80\Delta - c_d; (c_d = 0) \end{cases}$$

O portfolio será isento de risco se escolhermos um Δ que satisfaça a igualdade:

$$120D - 10 = 80D \therefore 40D = 10 \therefore D = 0,25$$

Portanto, o valor do portfólio no final do período será igual a 20 ($0,25 \times 120 - 10 = 0,25 \times 80$). Ao igualar o retorno obtido com a taxa livre de risco encontra-se o valor da opção. Novamente, o valor da opção não dependerá do retorno esperado da ação.

Se continuar este processo, expandindo a árvore várias vezes, chega-se ao modelo binomial. Depois de n etapas, S pode ter os seguintes valores : $u^n S, du^{n-1} S, \dots, d^n S$. Para achar o valor da opção, basta retroceder na árvore construída, formando 2^{n-1} portfólios livres de risco em cada etapa. Cox, Ross e Rubinstein (1979) chegam a seguinte fórmula:

$$c = S\Phi[a; n; p'] - Kr^{-n}\Phi[a; n, p]$$

onde

$\Phi[a; n; p']$ é a função de distribuição binomial complementar

$$p = \frac{r-d}{u-d} \quad ; \quad p' = (u/r)p$$

a : menor inteiro > 0 maior que $\log(X/Sd^n) / \log(u/d)$. Se $a > n$, $c = 0$.

A expressão encontrada é análoga à fórmula de Black e Scholes. Cox, Ross e Rubinstein (1979) mostram que, quando levada ao $\lim n \rightarrow \infty$, esta expressão converge para a solução da equação de Black e Scholes. Contudo, para se chegar a este resultado, u e d devem permanecer constantes ao longo do tempo.

4.1.4.3. Modelo do Salto (“Pure Jump Model”)

Outro modelo proposto por Cox, Ross e Rubinstein (1979) permite que o preço da ação varie de forma descontínua. Normalmente, o preço do ativo mudará de maneira contínua, mas existirá uma pequena probabilidade de, repentinamente, o preço dar um salto. Portanto, em cada pequeno intervalo de tempo, Δt , o preço da ação pode se mover de S para Su , por exemplo, com uma pequena probabilidade $\lambda \Delta t$ e para $Se^{-w\Delta t}$ com uma probabilidade $1 - \lambda \Delta t$.

Se levarmos este movimento ao limite $\Delta t \rightarrow 0$, saltos ocorrerão de acordo com um processo de Poisson a uma taxa λ . O preço da ação seguirá uma distribuição log-Poisson e o preço da opção de compra será dado por:

$$c = S\Psi(x; y) - Ke^{-r(T-t)}\Psi\left(x, \frac{y}{u}\right)$$

onde

$$\Psi(a; b) = \sum_{i=a}^{\infty} \frac{e^{-b} b^i}{i!} \quad ; \quad y = \frac{(r+w)(T-t)u}{u-1}$$

e x corresponde ao menor inteiro ≥ 0 que seja maior que $\frac{\ln(K/S) + w(T-t)}{\ln u}$

4.1.4.4. Modelo do Salto-Difusão

Merton (1976) sugeriu um modelo em que o preço da ação exhibe saltos superpostos ao movimento browniano:

$$\frac{dS}{S} = (m - Ik)dt + Sdz + dq$$

onde

m : retorno esperado do ativo

I : taxa na qual os saltos ocorrem

k : salto médio referido como aumento proporcional no preço do ativo

Para chegar a sua fórmula, Merton assume que o componente salto representa um risco não sistemático, podendo ser eliminado num portfólio diversificado e que portanto não é precificado na economia. Se o logaritmo do tamanho proporcional do salto tiver uma distribuição normal com desvio padrão d , o preço da opção de compra européia é:

$$c = \sum_{n=0}^{\infty} \frac{e^{-I't} (I't)^n}{n!} f_n$$

onde

$$t = T - t \quad ; \quad I' = I(1 + k)$$

f_n : c de Black e Scholes quando a taxa de variância instantânea é

$$s^2 + \frac{nd^2}{t}$$

e a taxa livre de risco igual a

$$r - Ik + \frac{ng}{t}; \quad g = \ln(1 + k)$$

4.1.4.5. Modelo de Volatilidade Estocástica

Neste modelo, Hull e White (1987) partem das seguintes equações:

$$dS = fS dt + sS dw$$

$$dV = mV dt + xV dz$$

$$V = s^2$$

dz e dw são processos de Wiener que têm correlação r .

Há uma dificuldade em encontrar a solução para este problema devido a inexistência de um ativo cujo preço seja perfeitamente correlacionado com s^2 . Sem tal ativo, não se consegue formar um portfolio livre de risco.

Hull e White (1987) conseguem encontrar uma solução para a equação diferencial que a opção deve satisfazer quando o preço da ação é instantaneamente não-correlacionado com a volatilidade. Também examinam soluções numéricas para os casos em que a volatilidade é correlacionada com o preço da ação.

Quando a volatilidade é estocástica mas não correlacionada com o preço do ativo, o preço de uma opção de compra européia é o preço de B-S integrado sobre a distribuição da taxa de variância média durante a vida da opção:

$$\int c(\bar{V})g(\bar{V})d\bar{V}$$

onde \bar{V} é o valor médio da taxa de variância, c é o preço de Black e Scholes e g é a função densidade de probabilidade de \bar{V} num mundo neutro ao risco.

Quando não existe correlação, ou, quando a correlação é próxima de zero, o preço da opção é menor que o de Black e Scholes quando a opção está perto ou no-dinheiro (at the money), e maior quando muito fora ou dentro-do-dinheiro. Os preços de exercícios, para os quais as opções são superestimadas por Black e Scholes, se encontram numa faixa de dez por cento do preço do ativo. Este efeito é aumentado, à medida que o prazo de vencimento aumenta. Portanto, se a equação de Black e Scholes for utilizada para se obter a volatilidade implícita de uma opção, quanto maior for o prazo de vencimento, menor será a volatilidade implícita.

Quando a correlação é significativamente positiva, a fórmula de Black e Scholes tende a subestimar o preço das opções de compra fora-do-dinheiro (out-of-the-money). Isto pode ser melhor entendido se pensarmos que, quando o preço da ação aumenta, a

volatilidade também aumenta, aumentando desta forma a probabilidade de que a opção seja exercida.

Quando a correlação é significativamente negativa, a fórmula de Black e Scholes tende a superestimar as opções de compra fora-do-dinheiro. Quando o preço da ação sobe, diminui a volatilidade fazendo com que seja menos provável que a opção seja exercida.

Apesar destes modelos sobrepor as limitações iniciais do modelo original de Black e Scholes, para utilizá-los tem-se que estimar diversos parâmetros o que aumenta o grau de incerteza. Black (1998) examina as premissas de sua fórmula original e sugere algumas modificações para melhorar a performance de seu modelo.

Britten-Jones e Neuberger (1998) analisam a validade da fórmula de Black e Scholes, deduzido a partir de premissas irrealistas, e chegam a conclusão que a fórmula original pode ser aplicada num mercado real, discreto, com custos caso se consiga prever corretamente a volatilidade dos preços negociados. Mas, a eficiência da fórmula depende do tamanho dos saltos apresentados pelos preços dos ativos. Se os saltos não forem superiores a 0,5%, a fórmula pode ser utilizada.

Uma outra alternativa para tentar-se sobrepor as limitações do modelo original de Black e Scholes é a utilização de modelos não paramétricos, apresentados a seguir.

4.2. Modelos Não-Paramétricos

Estes modelos não restringem os movimentos do preço da ação a nenhum modelo estrutural. Utilizam séries históricas para tentar obter o valor das opções. Campbell, Lo e MacKinlay (1997) reforçam o fato destes modelos, apesar de serem capazes de cobrir uma variedade de não-linearidades, precisarem de uma quantidade substancial de dados.

Dentro deste tipo de abordagem, tem-se os trabalhos de Stutzer (1996) utilizando valorização canônica e Hutchinson, Lo e Poggio(1994) que propõem o uso de redes neurais para precificar opções.

4.2.1. Valorização Canônica

Stutzer (1996) propõe uma forma não paramétrica chamada valorização canônica para estimar o preço das opções.

O processo de valorização canônica se dá em três etapas. Primeiro, dados históricos são computados para se construir a distribuição empírica de retornos. Depois, esta distribuição é transformada em uma estimativa de probabilidades ajustadas livres de risco (*martingale* equivalente). Finalmente, o valor do derivativo é estimado utilizando-se a distribuição de probabilidade ajustada para calcular seu valor esperado descontado.

O principal desafio do modelo é encontrar a distribuição livre de risco a partir da construída a partir de dados históricos. Stutzer (1996) utiliza um método sugerido por estatísticos e físicos para resolver este problema. Essa estimativa vem da resolução da minimização do critério de distância de Kullback-Leibler, I , das probabilidades positivas ajustadas, \mathbf{p}^* , para as probabilidades empíricas $\hat{\mathbf{p}}$

$$\hat{\mathbf{p}}^* = \arg \min_{\mathbf{p}^*(h) > 0, \sum_h \mathbf{p}^*(h) = 1} I(\mathbf{p}^*, \hat{\mathbf{p}}) = \sum_{h=1}^{H-T} \mathbf{p}^*(h) \log \frac{\mathbf{p}^*(h)}{\hat{\mathbf{p}}(h)}$$

onde T representa número de períodos da data de vencimento da opção e H o último período do qual se tem informação.

A solução do problema é a seguinte distribuição canônica de Gibbs:

$$\hat{p}^*(h) = \frac{\exp\left[\mathbf{g}^* \frac{R(-h)}{r^T}\right]}{\sum_h \exp\left[\mathbf{g}^* \frac{R(-h)}{r^T}\right]}, \quad h = 1, \dots, H - T.$$

onde R representa o retorno histórico, r a taxa livre de risco e \mathbf{g}^* o multiplicador de Lagrange que pode ser encontrado solucionado-se o seguinte problema

$$\mathbf{g}^* = \arg \min_g \sum_h \exp\left[\mathbf{g} \left(\frac{R(-h)}{r^T} - 1\right)\right]$$

Stutzer (1996) simula um ambiente onde a mudança do preço do ativo segue a as premissas do modelo de Black e Scholes, utiliza o método canônico para precificar opções e chega à conclusão que a precisão do método é comparável ao modelo de Black e Scholes, baseado nas volatilidades históricas, naquelas condições.

No entanto, Stutzer (1996) lembra que o método canônico pode não funcionar tão bem em casos onde as probabilidades corretas de *martingale* são radicalmente diferentes das probabilidades reais.

4.2.2. Precificação de Opções Utilizando Redes Neurais⁷

Hutchinson, Lo e Poggio(1994) propõem uma alternativa para a precificação de opções quando a dinâmica do ativo base não é conhecida ou quando a equação associada à condição de não arbitragem não pode ser resolvida analiticamente. Utilizando preços de opções e variáveis primárias que influenciam o preço das opções, “treinam” a rede neural que posteriormente se transforma na fórmula de precificação passando a gerar o preço das opções de acordo com os dados de entrada.

⁷ Outras aplicações das redes neurais no mercado financeiro podem ser encontradas em Refenes (1995) e Trippi (1993)

Entre as vantagens deste modelo pode-se citar a robustez em relação a erros de especificação que contaminam os outros modelos, a adaptabilidade das redes a responder a mudanças estruturais no processo de geração de dados e a flexibilidade para lidar com uma ampla gama de derivativos e dinâmicas do ativo base.

Sua principal desvantagem é o fato de precisar de uma substancial base de dados para se obter uma rede bem treinada. Portanto, não é apropriada para derivativos pouco negociados, ou recentemente criados. A não ser que estes derivativos possam ser reproduzidos pela combinação de outros preexistentes.

As redes neurais são examinadas a seguir.

4.2.2.1. Redes Neurais⁸

Uma rede neural pode ser descrita em termos de seus neurônios individuais, suas conexões (topologia) e regra de aprendizagem. Juntos estes três fatores constituem o modelo da rede neural.

Unidade Básica

Haykin (1999) faz um paralelo das redes neurais e o modo como o cérebro humano processa informações. Fazendo-se uma analogia com nosso sistema neurológico, pode-se dizer que sua unidade básica é o neurônio, constituído por um corpo central, dentritos para receber os dados de entrada e o axônio que transmite o sinal de resposta para os dentritos de outros neurônios. A junção entre um axônio e um dentrito é chamada de sinapse. Os sinais das sinapses são coletados e somados pelo neurônio. Se a força do sinal combinado excede um determinado patamar, o neurônio emite seu próprio sinal que é uma transformação do sinal original de entrada.

⁸ Explicações mais detalhadas sobre redes neurais podem ser encontradas em Haykin(1999)

Cada conexão tem um peso associado, w_i , de tal forma que o neurônio calcula sua própria saída, y , através da soma ponderada de suas entradas, x_i , gerando um nível de ativação, z , e passando através de uma função de transferência. A Figura 9 ilustra o processo.

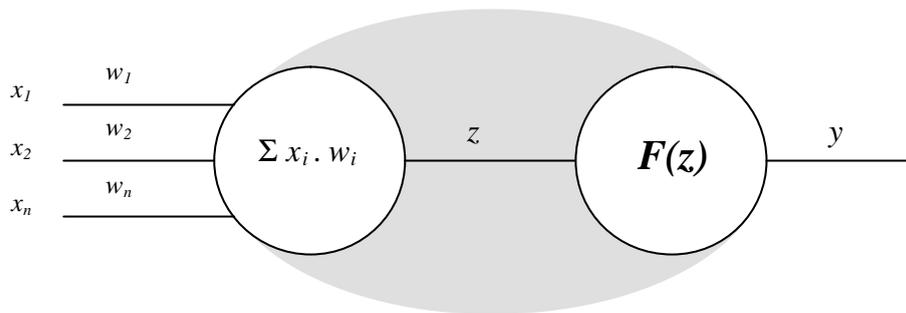


Figura 9: Representação do Neurônio

Então, cada neurônio produz uma resposta, que é o resultado da função de transferência utilizando-se as variáveis de entrada: $y = F\left(\sum x_i \cdot w_i\right)$

A função de transferência pode ser uma função degrau, sigmóide ou uma função pseudo-linear. Cada função de transferência introduz uma não-linearidade na rede enriquecendo sua capacidade representacional. A função de transferência mais utilizada é a função sigmóide. É definida por Haykin (1998) como uma função estritamente crescente que exibe um equilíbrio atrativo entre um comportamento linear e não linear.

$$F(z) = \frac{1}{1 + e^{-z}}$$

A Figura 10 ilustra a função.

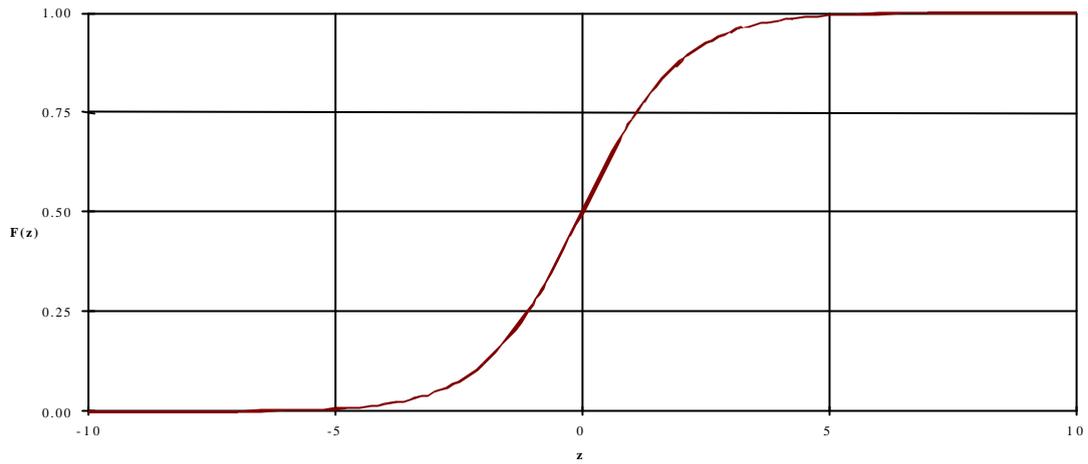


Figura 10: Função sigmóide

Estrutura

Uma rede neural é uma estrutura composta de neurônios primários interconectados como ilustra a Figura 11. Estes neurônios são tipicamente organizados em camadas, com cada neurônio em uma camada tendo sua conexão com cada neurônio da próxima camada. A maneira em que os neurônios estão estruturados está intimamente ligada com o algoritmo utilizado para treinar a rede.

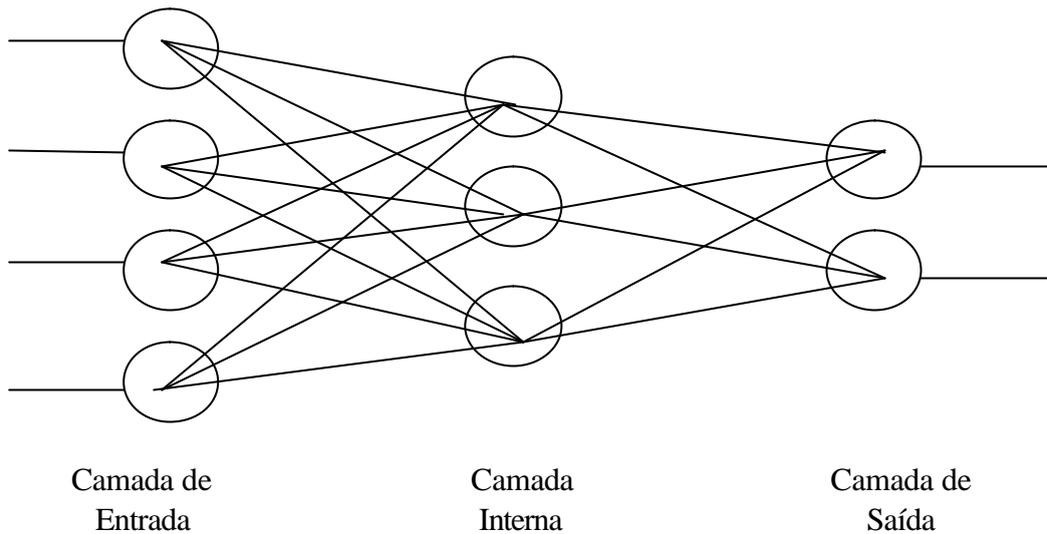


Figura 11: Representação de uma Rede Neural

O tipo de rede utilizado neste estudo é a multi-camada feed forward que tem uma ou mais camadas de neurônios. Cada camada recebe um vetor de entrada que é o vetor de saída da camada anterior. As camadas estão dispostas numa ordem linear, de tal forma que a entrada da primeira camada é a entrada externa, entrada da segunda é a saída da primeira e assim por diante, com cada camada recebendo seu vetor de entrada da camada anterior. A última camada é a de saída, e seu vetor de saída representa a saída da rede.

Os neurônios da camada interna tem a função de intervir entre os neurônios externos de entrada e saída. Através da adição de uma ou mais camadas internas a rede é capacitada a extrair estatísticas de alta ordem.

Sistema de “aprendizagem”

Haykin (1999) define a aprendizagem de uma rede neural como

“Aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação produzido pelo ambiente em que a rede está inserida. O tipo de aprendizagem é determinado pela maneira em que a mudança dos parâmetros ocorre.” (pág. 50)

Um sistema neural “aprende” através da mudança de suas conexões internas respondendo a estímulos externos, ou seja, a resposta da rede é adaptada alterando-se os pesos de suas conexões. O algoritmo de aprendizagem utiliza conjuntos de vetores de entrada e saída para treinar a rede. No chamado aprendizado supervisionado, especifica-se o padrão de saída desejado. Deste modo, um retorno de aprendizagem é fornecido para cada par de dados (entrada-saída) para ser aprendido pela rede. Assim, a diferença entre a saída desejada e a realizada é conhecida para todo conjunto de dados. O objetivo deste tipo de aprendizagem é eliminar as diferenças entre as saídas da rede e as desejadas minimizando uma função de custo ou maximizando uma função objetivo.

O algoritmo da retropropagação do erro (back propagation) é, provavelmente, o mais conhecido e utilizado algoritmo de aprendizagem supervisionado. Refenes (1995) afirma que a descoberta deste algoritmo, mais que a das redes, foi que proveu o ímpeto para a exploração prática das redes neurais. É um algoritmo de gradiente descendente, ou seja, ajusta os pesos da rede de tal forma que o erro global decresce continuamente. Um obstáculo desta técnica é que a rede pode se fixar num mínimo local e não encontrar o conjunto de parâmetros que determinam o mínimo global, isto é, o ponto ótimo.

Para iniciar o processo, os pesos são inicializados com valores aleatórios. O treinamento da rede se dá através da mudança sistemática dos pesos até que o sistema produza a saída desejada, dado um certo nível de tolerância. A força das conexões (pesos) é então ajustada na direção decrescente do erro. Uma medida bastante utilizada de erro é dada por:

$$E = \frac{1}{2} \sum_{j,c}^n (y_{j,c} - d_{j,c})^2$$

onde $y_{j,c}$ representa a resposta da unidade j para o input/output c , e $d_{j,c}$ a resposta desejada.

Portanto, a aprendizagem da rede é um processo de minimização da medida de erro. Depois que números aleatórios são escolhidos como pesos, usam-se as seguintes equações para ajustá-los:

$$\Delta w_{ij,k}^n = \mathbf{h} \cdot \mathbf{d}_{j,k} \cdot y_{i,k-1}$$

$$w_{ij,k}^{n+1} = w_{ij,k}^n + \Delta w_{ij,k}$$

$$\mathbf{d}_{j,k} = \left(\frac{dy}{dz} \right)_{i,k} \cdot (y_{i,k} - y_{i,k}^T)$$

para $i = 1, 2, \dots, I; j = 1, 2, \dots, J$ e $k = K$.

onde $w_{ij,k}^n$ é o peso conectando o neurônio i na camada $k-1$ ao neurônio j na camada k na iteração n ;

\mathbf{h} é coeficiente da taxa de treinamento (normalmente entre 0.01 e 1.0);

$y_{i,k}$ é a saída do neurônio i na camada k ;

$y_{i,k}^T$ é o valor desejado para $y_{i,k}$.

O algoritmo de retropropagação do erro consiste em duas passagens pela rede. Na primeira, de ida, o vetor de entrada é processado camada a camada até que, finalmente, a resposta da rede é produzida. Nesta passagem os parâmetros da rede permanecem fixos. Durante a segunda passagem, de volta, os parâmetros são ajustados de acordo com a regra de correção.

Note que o algoritmo acima só pode ser aplicado para a última camada de neurônios, uma vez que requer o conhecimento da saída desejada. Para as camadas internas não há saída desejada e, conseqüentemente, $\mathbf{d}_{j,k}$ tem que ser obtido de um outro modo. O algoritmo de retropropagação deriva seu nome do fato de propagar o valor de $\mathbf{d}_{j,k}$ para trás através da rede. Para as camadas internas, temos:

$$\mathbf{d}_{j,k} = \left(\frac{dy}{dz} \right)_{i,k} \cdot \left(\sum_j \mathbf{d}_{j,k} \cdot w_{ij,k} \right)$$

para $k = 1, 2, \dots, K-1$.

O coeficiente da taxa de treinamento controla a velocidade de aprendizagem. Quanto menor o valor do coeficiente, menores são as mudanças nos parâmetros da rede a cada iteração e mais suave é a trajetória de ajuste dos pesos. Esta melhora, no entanto, implica uma velocidade de aprendizagem mais lenta. Por outro lado, caso o coeficiente seja muito aumentado, a mudança nos parâmetros pode ficar instável, isto é, oscilatória.

Para se atingir um ajustamento mais rápido no algoritmo de retropropagação, evitando-se o problema de instabilidade, pode-se incluir uma constante de momento. Essa constante de momento é utilizada para ajudar a rede a seguir os estreitos vales, se existirem, na superfície de erro, evitando saltos rápidos de um lado para o outro. Isto se dá através da adição de um termo autoregressivo ao algoritmo. Assim, tem-se uma “memória” dos ajustes prévios

$$\Delta w_{ij,k}^{n+1} = \mathbf{h} \cdot \mathbf{d}_{j,k} \cdot y_{i,k-1} + \alpha \cdot \Delta w_{ij,k}^n$$

$$w_{ij,k}^{n+1} = w_{ij,k}^n + \Delta w_{ij,k}^{n+1}$$

onde α é a constante de momento.

A próxima seção examina as evidências empíricas da teoria de precificação de opções.

4.3. Evidências Empíricas

Uma forma de se escolher um método de precificação de opções é conduzir e observar estudos empíricos. No entanto, qualquer hipótese estatística sobre modelos de precificação de opções é uma hipótese conjunta de que a fórmula de precificação está correta e que o mercado é eficiente. Outra dificuldade está relacionada com a volatilidade que é uma variável que não pode ser observada. Pode-se estimá-la de várias maneiras: através de valores históricos, de valores implícitos nos prêmios das opções ou por outros métodos. Também deve-se ter cuidado para que os preços dos ativos primários e das opções estejam sincronizados para a realização de testes.

Black e Scholes (1973) assim como Galai (1977) testaram se era possível obter retornos superiores à taxa livre de risco comprando opções subvalorizadas pelo mercado em relação ao preço teórico, vendendo opções e mantendo uma posição delta-neutro. Ambos os estudos, na ausência de custos transacionais, obtiveram retornos superiores à taxa livre de risco.

Chiras e Manaster (1978) compararam a volatilidade ponderada implícita pelos preços das opções com a volatilidade histórica utilizando dados da *Chicago Board Options Exchange*. Chegaram à conclusão que o método implícito fornecia uma melhor previsão da volatilidade durante a vida da opção. Contudo, não há um consenso sobre o assunto. Suportando o modelo de B-S, obtiveram resultados positivos ao testar uma estratégia de comprar opções com baixa volatilidade implícita e vender as de alta volatilidade.

MacBeth e Merville (1979), utilizando opções de ações em 1976, chegaram à conclusão que as volatilidades implícitas tendem a ser relativamente altas para as opções

dentro-do-dinheiro e relativamente baixas para opções fora-do-dinheiro. Estes resultados tornam-se mais pronunciados quando aumenta-se o prazo das opções e o grau em que a opção se encontra dentro ou fora do dinheiro.

Rubinstein (1985) realizou um estudo similar ao de MacBeth e Merville comparando volatilidades implícitas de opções, entre 23/08/1976 a 21/10/1977, que diferiam apenas na data de exercício ou no preço de exercício. Dividiram o período de testes em dois subperíodos, 23/8/76 a 21/10/77 e 22/10/77 a 31/8/78. No primeiro intervalo, os resultados são consistentes com os de MacBeth e Merville (1979). No entanto, resultados opostos foram encontrados no segundo intervalo, ou seja, as volatilidades implícitas eram relativamente altas para opções fora-do-dinheiro e relativamente baixas para opções dentro do dinheiro. Durante todo o período, as opções de curto prazo tinham volatilidades implícitas significativamente mais altas. Estes resultados indicam que houve uma mudança na correlação entre o ativo e a volatilidade.

Relaxando algumas premissas restritivas do modelo de Black e Scholes, Bakshi, Cao e Chen (1997) derivam modelos genéricos que admitem volatilidade estocástica, taxa de juros estocástica e “saltos” aleatórios. Utilizando opções do S&P 500, examinaram vários modelos sob três perspectivas:

1. Consistência interna dos parâmetros implícitos com as séries temporais: todos os modelos apresentam inconsistências, sendo o modelo de salto-difusão e volatilidade estocástica o menos inconsistente, e o Black e Scholes o mais inconsistente.

2. Precificação fora da amostra: Black e Scholes apresenta os maiores erros, seguido do modelo de volatilidade estocástica. O salto-difusão apresenta os menores erros. A inclusão de taxa de juros estocástica melhora o ajuste para opções de longo prazo.

3. Performance de hedge: Dois tipos de estratégias de hedging foram utilizados para avaliar a eficiência dos modelos. A primeira estratégia consistia no hedge utilizando o ativo primário como único instrumento de hedging. Neste tipo de hedge, o modelo de volatilidade estocástica tem o melhor desempenho seguido do modelo de volatilidade estocástica com saltos aleatórios. Na segunda estratégia, a posição líquida era completamente imunizada (localmente) através do uso de um número de instrumentos igual ao número de fontes de riscos. Os erros de hedge para os modelos de volatilidade estocástica, volatilidade estocástica - taxa de juros estocástica e volatilidade estocástica-saltos aleatórios foram de 50 por cento a 65 por cento mais baixos que o de B-S. A mudança da frequência de hedge afetava significativamente os erros provenientes do modelo de B-S e marginalmente os outros modelos.

Bakshi, Cao e Chen (1997) observaram que, em geral, independentemente da medida de performance adotada, a incorporação da volatilidade estocástica é o fator mais importante no aprimoramento da fórmula de Black e Scholes. De acordo com os resultados, adicionando-se a característica dos saltos aleatórios melhora ainda mais a precificação do modelo de volatilidade estocástica, especialmente, as de curto prazo; enquanto que os modelos de taxa de juros estocástica ajudam a precificação das de longo prazo.

Assumindo que a volatilidade dos retornos de um ativo é uma função determinística do preço do ativo e tempo, Dumas, Fleming e Whaley (1998) desenvolveram um modelo de valoração de opções utilizando uma função determinística de volatilidade capaz de ajustar-se aos preços das opções negociadas. A idéia de se tentar ajustar os preços de

opções por uma função determinística é muito similar à construção de uma árvore binomial ou trinomial que se ajusta aos preços das opções.

Tentando superar os desvios do modelo de B-S em descrever os preços das opções utilizando uma volatilidade constante, Dumas et alli testam a hipótese da volatilidade ser uma função determinística do preço do ativo e do tempo. Para testar o modelo, verificam se a função de volatilidade implícita “hoje” é a mesma que está embutida nos preços de opções “amanhã”. Utilizando dados de opções do S&P 500 de Junho de 1988 a Dezembro de 1993, examinaram a performance preditiva e de hedging de seu modelo que não foi superior ao modelo de Black-Scholes ajustado às volatilidades implícitas aos preços de exercícios e prazos de vencimentos.

Cada função de volatilidade era estimada através da minimização da soma dos quadrados dos erros entre os preços reportados e os calculados pelo modelo. Apesar da alta flexibilidade em especificar a função de volatilidade, que faz com que sempre seja possível descrever exatamente a estrutura dos preços das opções, os resultados da pesquisa mostram que quanto mais simples o modelo melhor. Quando as funções de volatilidade foram utilizadas para precificar as opções uma semana mais tarde, os erros cresciam à medida que o modelo se tornava menos parcimonioso. Especificamente, modelos que incluem um parâmetro temporal têm os piores desempenhos, indicando que a variável tempo é uma importante causa de *overfitting*, ou seja, apesar de melhorar o ajuste do modelo aos dados prejudica a capacidade de generalização.

5. Método de Estudo

Para tentar descobrir a relação entre os fatores determinantes do preço das opções, treinou-se uma rede neural com dados de mercado. Os dados utilizados são descritos na seção 5.2..

Após treinada, utilizou-se a rede para a tomada de decisão de compra ou venda das opções comparando-se os preços gerados pelo modelo e os praticados no mercado.

Para avaliar a performance da rede fez-se uma análise dos resultados de uma estratégia de *hedge* dinâmico “*delta-neutro*”.

5.1. A Rede

Para construir, treinar e simular a rede neural, utilizou-se o simulador de redes neurais do software MATLAB. Não foi utilizado nenhum critério específico como adição ou remoção de neurônios para se determinar a estrutura da rede.

A rede utilizada foi uma rede multi-camada feed forward⁹ com duas camadas escondidas. Cada camada escondida possuía 5 neurônios cada e utilizava a função sigmóide como função de transferência. A camada de saída, por sua vez, utilizava uma função linear. A rede pode ser visualizada na Figura 12.

O algoritmo utilizado para treinar a rede foi um algoritmo de r-regularização automática do MATLAB(‘trainbr’). Os algoritmos de regularização tentam diminuir o problema de *overfitting*, produzindo redes que generalizam bem. Através da utilização destes algoritmos, diminui-se a dificuldade em se determinar a arquitetura ótima da rede.

⁹ Apresentada na seção 4.2.2.1.

O problema de *overfitting* ocorre quando o erro do conjunto de treinamento é levado a um valor bem pequeno, mas quando novos dados são apresentados à rede o erro é grande. Isto indica que a rede memorizou a amostra de treinamento, mas não aprendeu a generalizar para novas situações.

Um método para melhorar o poder de generalização de uma rede é utilizar uma arquitetura de rede grande que não seja maior que a necessária para se adequar aos dados. Quanto maior for a rede, mais complexas são as funções que ela pode criar e maior a probabilidade de sobrepor os dados sacrificando seu poder de generalização. A dificuldade está em saber de antemão quão larga deve ser a rede neural para uma aplicação específica.

Uma outra forma de se melhorar o poder de generalização de uma rede é através da utilização de técnicas de regularização. Alterando-se a função de performance da rede através da adição de um termo de soma quadrática dos parâmetros da rede, força-se a resposta a ser mais suave. Desta forma, a rede torna-se menos propensa a sobrepor os dados minimizando o problema de *overfitting*.

Antes de serem apresentados à rede, os dados passaram por uma etapa de pré-processamento. Os dados foram transformados de forma que ficassem numa faixa de valores $[-1,1]$ que possibilitasse uma melhor eficiência da rede.

5.2. Fatores Utilizados

Para treinar a rede foram utilizados os fatores do modelo de Black e Scholes: o preço da ação, o prazo de vencimento, o preço de exercício, a taxa de juros e volatilidade.

- *Preço da ação*: preço da ação da Telebrás no momento que foi negociada a opção.
- *Prazo de Vencimento*: número de dias úteis até a data de vencimento.

- *Preço de Exercício*: Preço de exercício das séries. Foram ajustados nas datas que a ação recebeu dividendos.
- *Taxa de juros*: Foi construída a curva diária de juros local com base nos preços de fechamento dos contratos futuros de juros da BM&F. A variável de entrada da rede é o fator de desconto efetivo até a data de vencimento da opção.
- *Volatilidade*: Como indicador da volatilidade da ação, foi utilizado o desvio-padrão dos preços das ações da Telebrás negociadas nos dois últimos dias (“*intra-day*”).

Os dados utilizados para estimar as variáveis da rede foram:

- Série histórica dos preços das ações da Telebrás de dezembro de 1997 a abril de 1999 negociadas na BOVESPA¹⁰
- Série histórica dos preços das opções de compra da Telebrás de dezembro de 1997 a abril de 1999, negociadas na BOVESPA, das seguintes séries: FEV98-130, ABR98-140, JUN98-140, AGO98-140, OUT98-90, DEZ98-90, FEV99-120 e ABR99-150
- Série histórica dos preços dos contratos futuros de juros (DI) de dezembro de 1997 a abril de 1999 negociados na BM&F¹¹
- Série histórica das taxas de CDI¹² registradas na CETIP¹³

Outros fatores, como volume negociado, *bid-ask spread*, número de oferta de compra e venda, variáveis binárias que representassem a existência de anúncio de informações relevantes, taxa de câmbio ou cotação do índice Dow Jones poderiam ter sido

¹⁰ Bolsa de Valores do Estado de São Paulo

¹¹ Bolsa de Mercadorias e Futuros

¹² Certificado de depósito interbancário

¹³ Central de Liquidação e Custódia de Títulos Privados

utilizados. De fato, talvez a performance de uma rede formada a partir destes fatores seja ainda melhor que a rede criada nesse estudo.

Mas, como pretendia-se comparar o desempenho da rede neural com o modelo de Black e Scholes optou-se por se utilizar as mesmas variáveis de entrada. A única diferença se refere à variável utilizada como indicador de volatilidade, pois pretendia-se utilizar uma variável que fosse independente do modelo de Black e Scholes para que não houvesse nenhuma interferência deste modelo no aprendizado da rede.

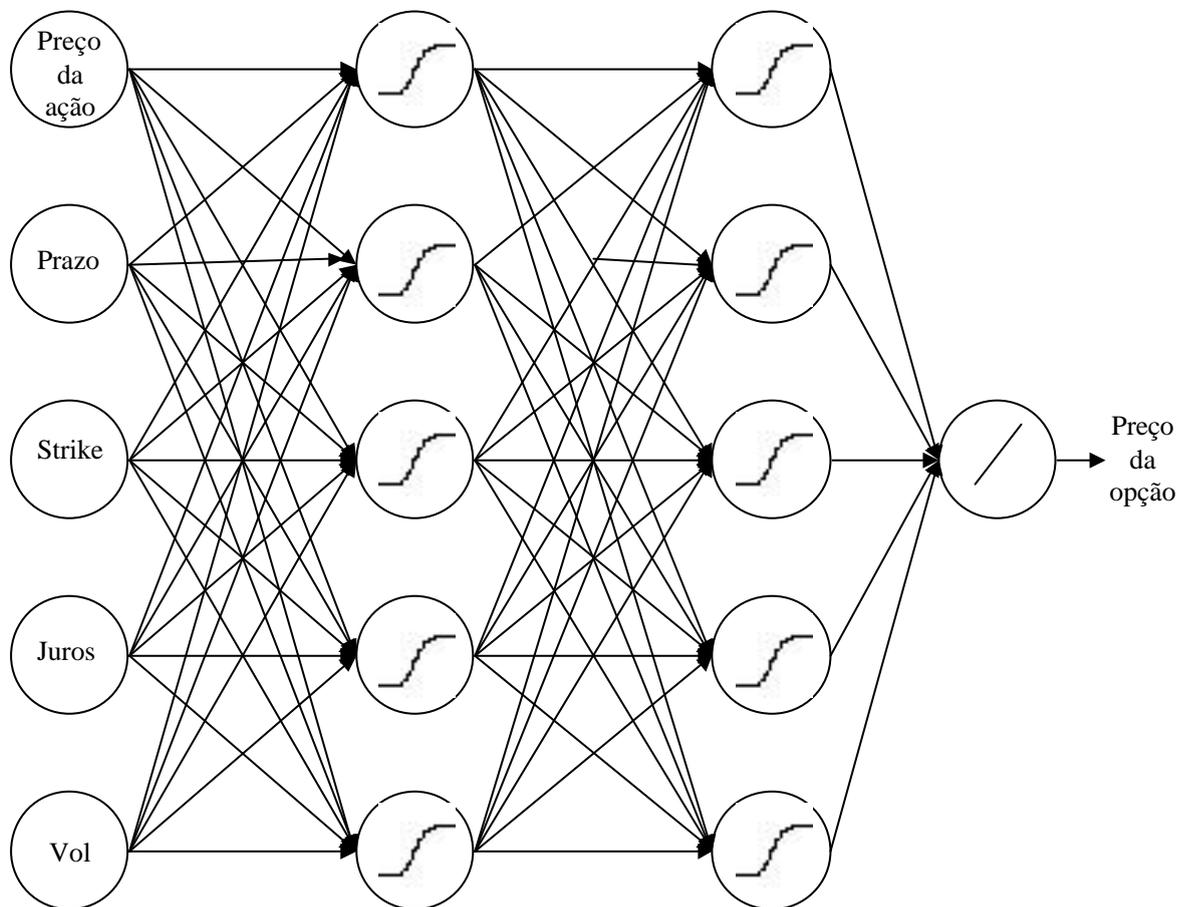


Figura 12: Rede utilizada - multi-camada feed forward com 2 camadas escondidas (5 neurônios utilizando a função sigmóide) e função linear na saída

5.3. Amostra

Para treinar e testar a rede neural, utilizou-se dados históricos relativos a oito séries de opções de compra de Telebrás entre dezembro/97 e abril/99.

Os dados foram divididos aleatoriamente em duas séries: treinamento e teste. A série de treinamento serviu para ajustar os parâmetros da rede. A série teste serviu para verificar a resposta da rede quando testada com dados que não haviam sido apresentados à rede na fase de adaptação de seus parâmetros.

Durante o período analisado, o preço da ação da Telebrás apresentou uma flutuação bastante intensa. O comportamento dos preços da Telebrás durante o período estudado pode ser observado na Figura 13.

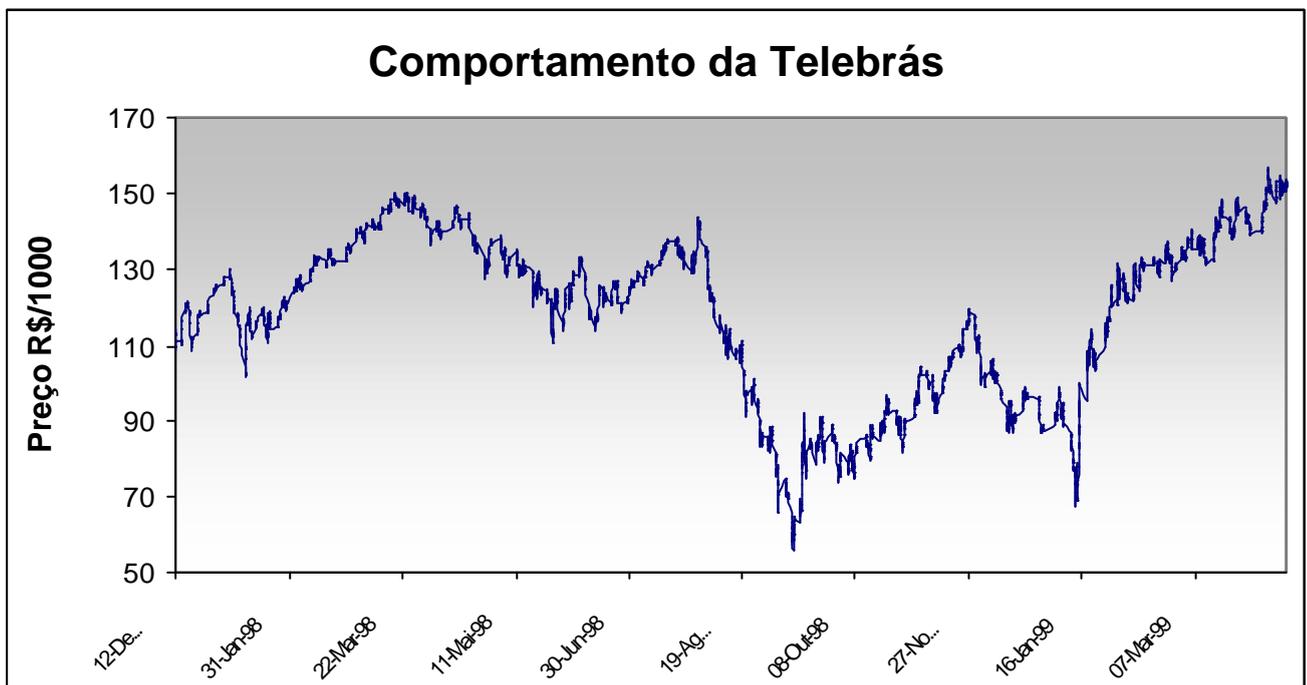


Figura 13: Comportamento da Telebrás durante o período estudado

5.4. Hedge Dinâmico

Para testar a performance da rede neural comparou-se os resultados de uma estratégia de *hedge* dinâmico *delta-neutro* com os resultados que seriam obtidos com o modelo de Black e Scholes. Além disso, tomou-se a saída da rede para tomada de decisão. Se o preço sugerido pela rede fosse maior que o praticado pelo mercado, a opção era comprada. Caso contrário, a opção era vendida.

A estratégia *delta-neutro* consiste no posicionamento em opções e ações de tal forma que o portfólio fique imune a pequenas oscilações do preço da ação. O número de ações a ser vendido ou comprado é determinado pelo *delta* da opção (sensibilidade do preço da ação à variação do preço da opção). Este procedimento já foi ilustrado na seção 4.1.2.2.

Becker, Lemgruber e Felício (1991), utilizando o mesmo princípio, reproduzem um opção de venda ilustrando como pode ser feito um seguro dinâmico de portfólio quando não há um mercado de opções disponível. Para comprovar empiricamente a eficácia da formação da opção sintética, realizam, com sucesso, uma simulação com dados históricos do mercado brasileiro de ouro de junho de 1988 a maio de 1990.

No modelo de Black e Scholes, o *delta* pode ser obtido derivando-se a solução em relação ao preço do ativo primário analiticamente. Como a função da rede neural é bastante complexa, fez-se uma aproximação numérica para se determinar a primeira derivada da rede em relação ao preço da ação. Portanto, para calcular o *delta* da rede calcula-se novamente o preço da opção utilizando os mesmos parâmetros utilizados anteriormente com um novo preço da ação (preço anterior + 0,10), subtrai-se o preço anterior e divide-se o resultado pelo acréscimo dado ao preço da ação.

A diferença entre os deltas sucessivos indica a quantidade marginal de ações a serem compradas/vendidas para que o portfólio continue delta-neutro.

Apesar da fórmula de Black e Scholes ter seu *delta* amplamente conhecido, precisa-se calcular a volatilidade implícita da opção para determiná-lo. A volatilidade implícita das opções foi determinada utilizando-se o método numérico de Newton-Raphson¹⁴.

Para verificar a performance dos dois modelos, simulou-se o *hedge* dinâmico das opções através do rebalanceamento dos portfólios conforme cada modelo. Ao final de cada dia, como ocorre no mercado, o saldo final era remunerado/cobrado em um CDI:

$$\text{Saldo Credor}_{d-1} : \text{Saldo Inicial}_d = \text{Saldo Credor}_{d-1} \times (1 + \text{CDI})$$

$$\text{Saldo Devedor}_{d-1} : \text{Saldo Devedor}_d = \text{Saldo Devedor}_{d-1} \times (1 + \text{CDI})$$

Seguiu-se com a estratégia até 4 dias úteis antes do vencimento da opção, quando o portfólio era liquidado.

¹⁴ O código utilizado no Excel pode ser visualizado no Apêndice 1

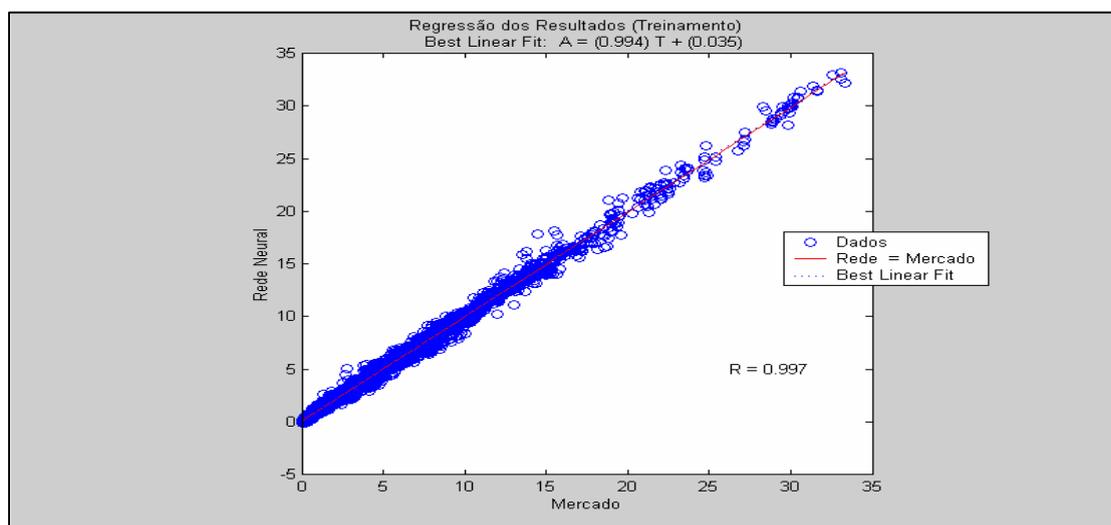
6. Resultados

Como a função da rede é precificar as opções, um dos primeiros resultados a ser analisado é o coeficiente de determinação, R^2 , dos valores estimados, comparados com os valores de mercado. Os resultados mostram uma grande proximidade entre os resultados esperados e os obtidos pela rede neural.

As Figuras 14 e 15 mostram a regressão e o coeficiente de correlação entre as respostas da rede e os valores de mercado para as fases de treinamento e teste respectivamente.

A análise dos resultados da amostra de teste mostra que a rede não só se adaptou à série de treinamento, mas também atingiu um grande poder de generalização pois, quando processou a série de testes, seus parâmetros já estavam determinados.

Figura 14: Regressão dos resultados da rede contra os valores de mercado para a amostra de



treinamento

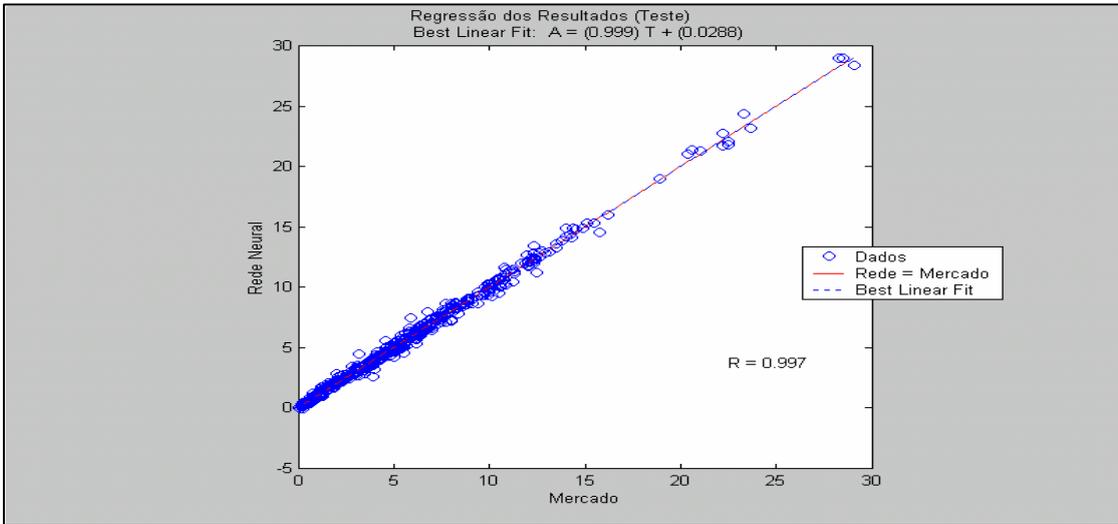


Figura 15: Regressão dos resultados da rede contra os valores de mercado para a amostra de teste

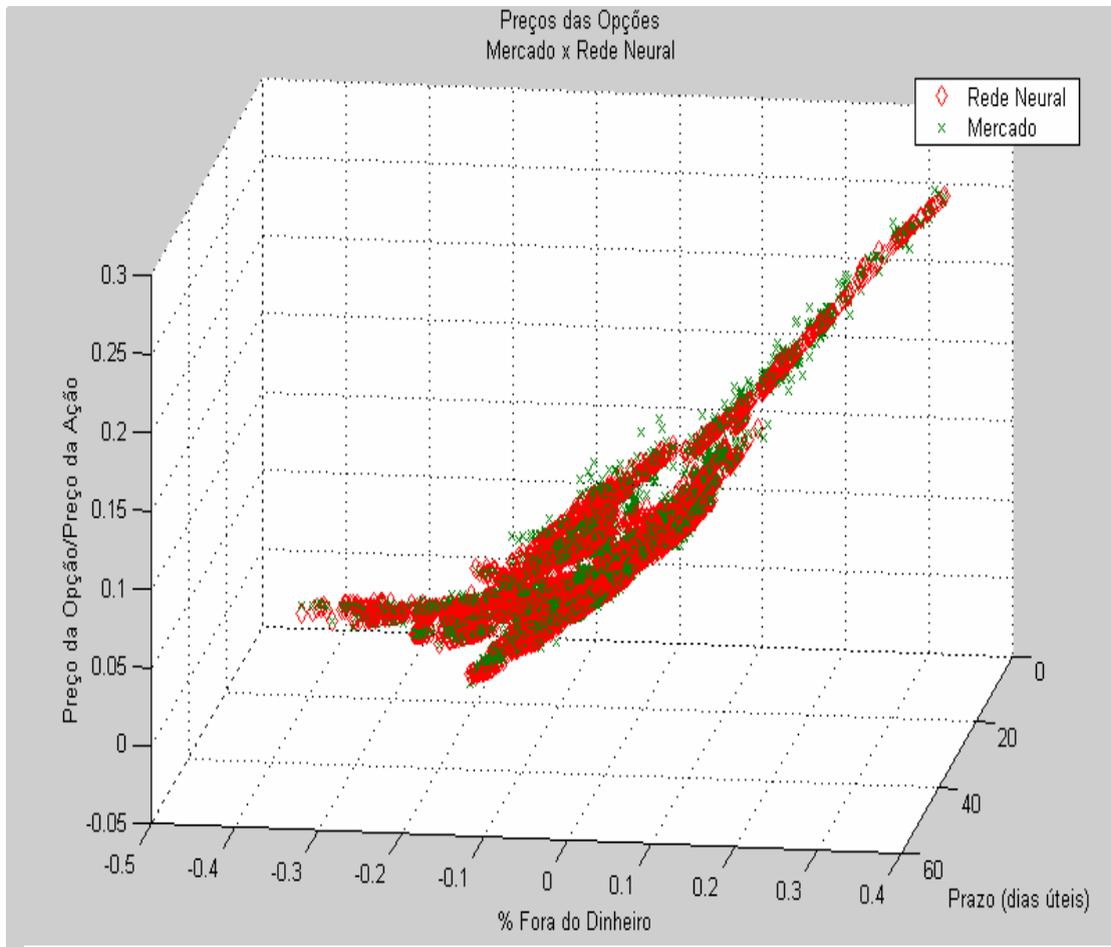


Figure 16: Preço das opções de compra divididos pelo preço da ação da amostra utilizada

A Figura 16 mostra os preços de mercado das opções de compra divididos pelo preço da ação junto com os preços previstos pela Rede Neural conforme o prazo de vencimento e o distanciamento do preço de exercício. Pode-se notar que o algoritmo de treinamento da rede tentou evitar o problema de ‘*overfitting*’ favorecendo a suavização da superfície ao invés de sobrepor todos os dados.

Novamente, pode-se notar o grande poder de adaptabilidade da rede. No entanto, a rede não foi parametrizada para ter como variável de saída valores positivos. Pode-se verificar que a rede produziu alguns valores de preço negativos para opções muito fora do dinheiro. Vale lembrar, como mostra a Figura 16, que a maioria das opções analisadas não se encontram nessas condições prejudicando a parametrização da rede para lidar com estas situações.

Pode-se comparar, também, os resultados da rede neural com o modelo de Black e Scholes através de uma simples inspeção visual das superfícies de preço e delta dos respectivos modelos. As Figuras 17 e 18 mostram a superfície de preços do modelo de Black e Scholes e da rede neural, utilizando a volatilidade média da amostra.

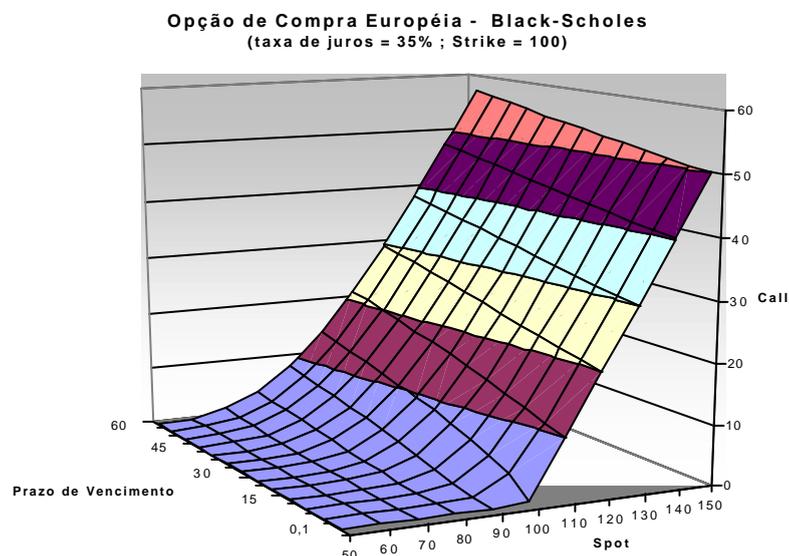


Figura 17: Preços das opções de compra, utilizando-se a volatilidade média da amostra, segundo Black e Scholes

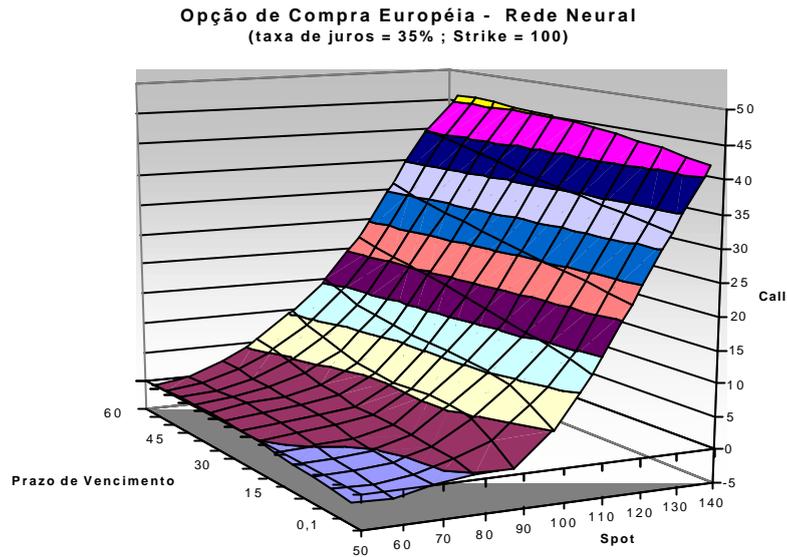


Figura 18: Preços das opções de compra, utilizando-se a volatilidade média da amostra, segundo o modelo da rede neural

Reforçando mais uma vez o poder de generalização da rede neural, a Figura 18 apresenta uma superfície relativamente uniforme devido a utilização da técnica de regularização.

As Figuras 19 e 20 mostram a superfície de delta do modelo de Black e Scholes e da rede neural utilizando a volatilidade média da amostra.

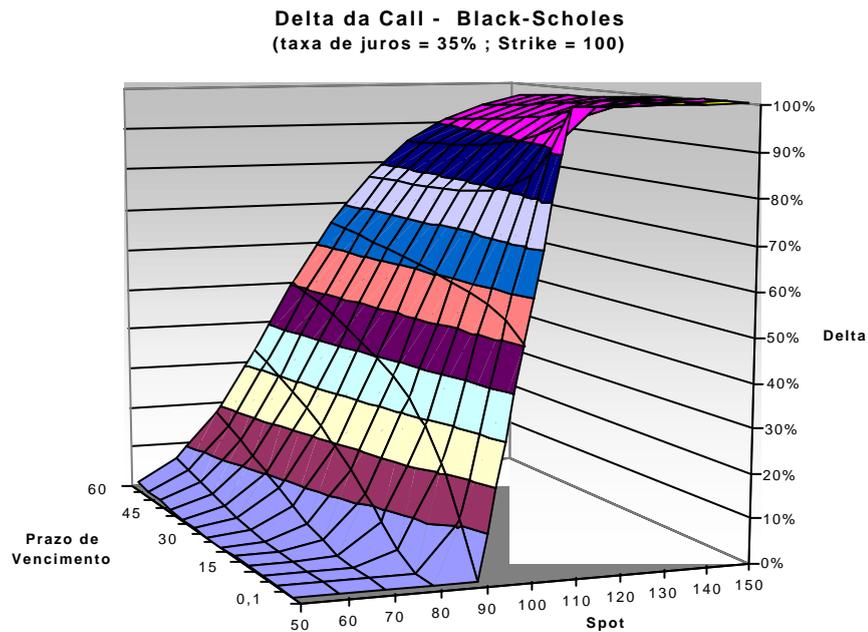


Figura 19: Delta das opções de compra de acordo com o modelo de Black e Scholes

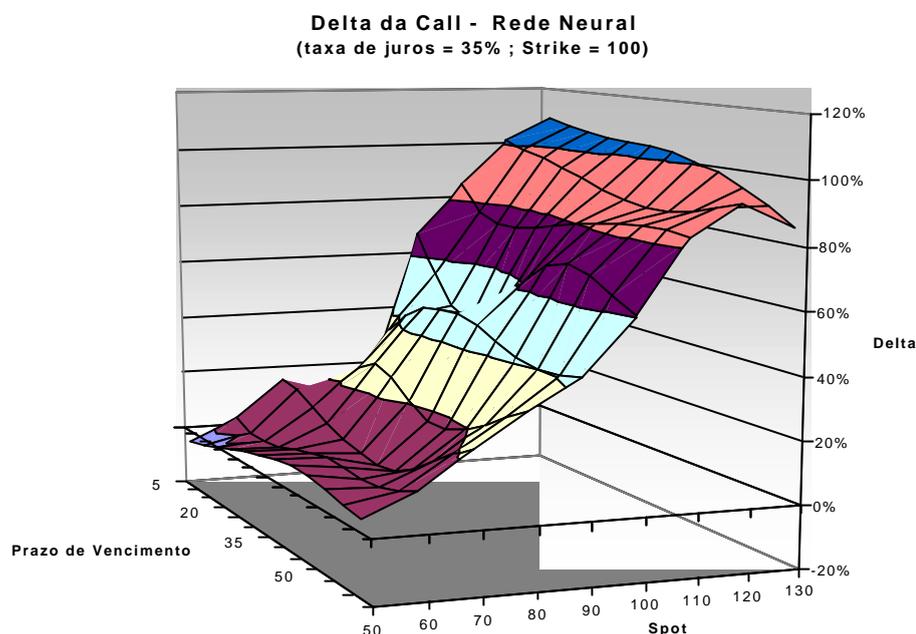


Figura 20: Delta da opção de compra de acordo com o modelo da Rede Neural

As Figuras 19 e 20, por sua vez, já apresentam uma diferença de superfície significativa o que refletirá nos resultados de hedge dinâmico das opções, uma vez que esses resultados dependem do delta estimado.

Uma medida mais significativa da performance de uma fórmula de precificação de opções e, que responde ao problema analisado por esse estudo, se dá através da mensuração do erro do portfólio de uma estratégia de hedge dinâmico utilizando-se a fórmula em questão para calcular a proteção necessária ou '*delta-hedge*'.

Os resultados da Tabela 1 indicam uma performance de hedge dinâmico superior da rede neural em relação ao modelo de Black e Scholes. Em média, o hedge dinâmico da rede neural apresentou um resultado superior R\$ 0,06 por opção ao modelo de B-S. Este desempenho foi verificada não só para a amostra integral mas também para a amostra de

teste isoladamente onde o resultado do hedge dinâmico da rede superou o do modelo de B-S em R\$0,17 por opção.

De acordo com os resultados obtidos pode-se rejeitar H_0 em favor de H_1 , e afirmar que os resultados de hedge dinâmico da rede neural são superiores ao do modelo de Black e Scholes com um grau de significância menor que 1%¹⁵. A Tabela 1 mostra o resumo dos resultados obtidos. A Tabela 2 e a Figura 21 mostram a dispersão da amostra em relação ao distanciamento do preço de exercício.

Delta hedge da rede neural e do modelo de Black e Scholes (R\$)

	RN	BS	RN - BS	RN	BS	RN - BS
		Amostra Integral			< -10%	
Mínimo	-7,73	-5,80	-4,14	-7,71	-5,80	-4,14
Máximo	8,74	6,89	4,08	8,30	6,40	4,08
Média	0,26	0,21	0,06	0,13	0,14	-0,01
Desvio	2,17	1,88	1,31	1,58	1,44	1,56
N	6340	6340		1235		
Teste T	0,025%			42,681%		
		> -10 & < -5%			> -5% & < 0%	
Mínimo	-7,73	-5,73	-3,27	-5,73	-4,31	-2,96
Máximo	8,74	6,89	3,16	5,21	4,39	2,98
Média	0,23	0,15	0,09	0,34	0,27	0,07
Desvio	2,97	2,44	1,27	2,11	1,74	1,34
N	1787			1810		
Teste T	0,189%			1,187%	2,374%	
		>0 & < 5%			>5%	
Mínimo	-2,81	-3,10	-3,51	-2,52	-3,74	0,05
Máximo	5,01	3,80	3,34	4,76	3,80	0,35
Média	0,25	0,20	0,04	0,33	0,08	0,12
Desvio	1,62	1,57	1,14	1,56	1,00	0,08
N	901			607		
Teste T	12,218%			0,610%		
		Teste				
Mínimo	-7,35	-5,54	-4,10			
Máximo	7,95	5,87	4,06			
Média	0,35	0,18	0,17			
Desvio	2,27	1,91	1,27			
N	500					
Teste T	0,168%					

Tabela1: Diferença entre o resultado de hedge dinâmico da Rede Neural (RN) e do modelo de Black e Scholes (BS) medido em Reais e segmentados pelo distanciamento do preço de exercício

¹⁵ Pode não haver independência entre as séries, o que comprometeria a significância do teste. Para sobrepor este problema, pode-se realizar simulações de Monte Carlo. No entanto, perde-se a aplicabilidade prática uma vez que não estaria se trabalhando com dados de mercado.

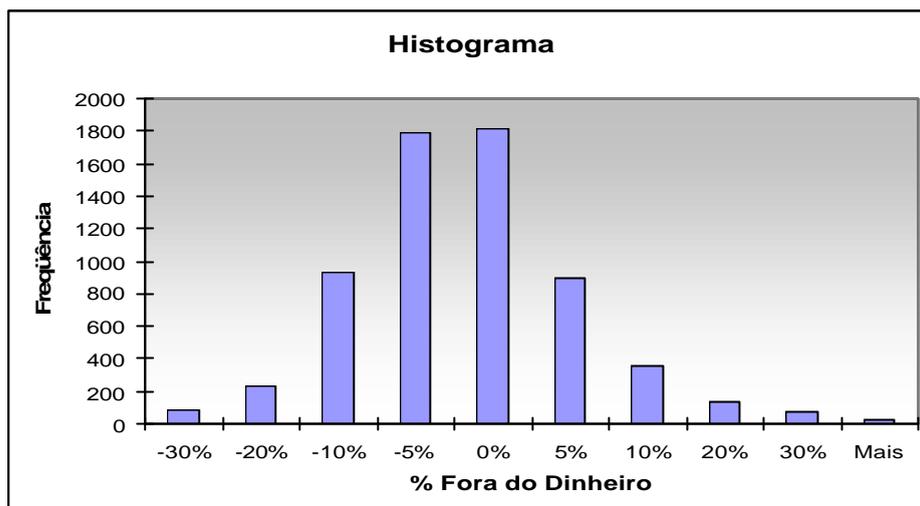


Figura 21: Histograma da amostra em relação ao distanciamento do preço de exercício

<i>% Fora do Dinheiro</i>	<i>Freqüência</i>	<i>% cumulativo</i>
-30%	83	1,31%
-20%	228	4,91%
-10%	924	19,48%
-5%	1787	47,67%
0%	1810	76,21%
5%	901	90,43%
10%	360	96,10%
20%	139	98,30%
30%	78	99,53%
Mais	30	100,00%

Tabela 2: Dispersão da amostra em relação ao distanciamento do preço de exercício.

Apesar dos resultados de cobertura dinâmica favorecerem fortemente as redes neurais, não se pode esquecer a falta de precisão desse modelo que chegou a preços negativos para as opções. Este problema pode ser contornado através da especificação da função de saída. No entanto, este procedimento não garantiria a precisão para valores positivos.

O impacto dos custos de transação não foram examinados neste estudo, uma vez que o objetivo era comparar a performance de duas abordagens. Como o hedge dinâmico era realizado na mesma frequência, o impacto dos custos de transação não alteraria os resultados finais.

7. Conclusões

Este trabalho apresentou uma comparação da performance de cobertura dinâmica entre o modelo de Black e Scholes ajustado com a volatilidade implícita de mercado e um modelo não paramétrico, redes neurais, formado pelas mesmas variáveis da fórmula de Black e Scholes com exceção do indicador de volatilidade utilizado pela rede.

A principal diferença entre a criação de uma rede neural e sua contraparte natural, um processador de informação clássico para classificação de padrões, é dada por Haykin (1999):

“No último caso, geralmente procedemos formulando, primeiro, um modelo matemático de fenômenos observados, validando o modelo com dados reais e então construindo com base no modelo. Contrastando com esta abordagem, a criação de uma rede neural é baseada diretamente em dados reais, sendo permitido ao conjunto de informações falarem por si. Portanto, a rede neural não só provém o modelo implícito nas condições do ambiente no qual está envolvida mas também executa a função de processamento de informações de interesse” (pág. 25)

Apesar da grande popularidade do método de Black-Scholes entre os operadores de mercado, os resultados deste estudo sugerem que as redes neurais devem ser consideradas como uma alternativa para os operadores de opções já que apresentaram uma performance de cobertura dinâmica de risco superior ao modelo de Black e Scholes.

De acordo com os resultados obtidos, a performance de hedge dinâmico da rede neural foi superior a do modelo de Black e Scholes. Por outro lado, na amostra selecionada não consegue-se garantir a independência entre as séries. Uma simulação de Monte Carlo pode resolver este problema, mas perde-se a aplicabilidade prática¹⁶.

¹⁶ Hutchinson, Lo e Poggio (1994) utilizaram Monte Carlo para simular um ambiente onde as premissas do modelo de B-S eram válidas e treinaram uma rede neural que teve uma performance de hedge dinâmico superior ao modelo de B-S no ambiente simulado

Ainda não podemos afirmar que a abordagem das redes neurais é superior às abordagens paramétricas. Este estudo focou apenas o modelo de Black-Scholes e sua aplicação foi testada apenas nas opções de compra de Telebrás entre dezembro/97 e abril/99. Além disso, este método não-paramétrico demanda uma grande quantidade de dados para se obter uma rede bem treinada. Conseqüentemente, esta abordagem não é apropriada para derivativos ilíquidos ou para novos derivativos que não possam ser replicados por outros já existentes.

Com relação à precisão da rede neural na determinação do preço absoluto das opções, os resultados deste estudo indicam que o modelo desenvolvido nesse estudo não deve ser utilizado como ferramenta única para o processo de tomada de decisão de compra ou venda de opções. Como afirma Hutchinson, Lo e Poggio (1994), se a dinâmica do ativo primário for bem conhecida e uma expressão analítica estiver disponível para esta dinâmica, a abordagem paramétrica quase sempre terá uma precisão melhor que a rede neural na precificação e cobertura dinâmica dos riscos. Contudo, como estas condições nem sempre ocorrem, ainda há um grande espaço para a aplicação prática das redes neurais na precificação de opções.

Por outro lado, os resultados obtidos nesse estudo são encorajadores se for levado em conta que somente um tipo de estrutura de rede e algoritmo de aprendizagem foi utilizado. Importantes variáveis que podem ter um reflexo na volatilidade do preço da ação, como o volume negociado, não foram consideradas.

Portanto, pode-se dizer que o potencial das redes neurais ainda não foi totalmente explorado e que resultados mais promissores podem ser esperados. Estes resultados podem ser melhorados através da substituição ou aumento dos fatores de entrada da rede.

Fatores como retorno médio do Ibovespa, volatilidade média do Ibovespa, volume de ações negociados e outras medidas de mercado podem ajudar a explicar o comportamento do ativo primário e são facilmente incorporados à estrutura de uma rede neural.

Uma outra forma de aumentar o potencial do uso das redes neurais pode vir do avanço das pesquisas na especificação das estruturas da rede. Novos métodos para determinar a arquitetura da rede, a interconexão entre os neurônios ou novos algoritmos de aprendizagem podem alavancar o uso desta ferramenta.

Estudos posteriores poderiam investigar a utilização de alguns dos fatores já mencionados e a aplicabilidade desta técnica a outros ativos financeiros como a taxa de câmbio. Uma outra sugestão seria o aproveitamento da capacidade preditiva das redes neurais juntamente com um modelo paramétrico para a cobertura de risco dinâmico das opções, que poderia acontecer através da criação de uma rede neural prever a volatilidade futura de determinado ativo ou do próprio preço deste ativo.

Finalmente, tendo em vista as análises e discussão desenvolvidas neste trabalho, espera-se que os profissionais do mercado financeiro brasileiro se interessem pelo uso desta ferramenta não paramétrica. Assim, maiores serão as chances de aperfeiçoamento e oportunidades de construir uma rede neural cuja saída seja uma estimativa da volatilidade futura do ativo primário, que poderá servir de entrada para algum modelo paramétrico.

Apêndice 1: Determinação da Volatilidade Implícita¹⁷*Mktprice: preço da opção**Strike: preço de exercício**Tenor: prazo em dias úteis**Spot: preço da ação**Fator: taxa de juros efetiva**Vol: volatilidade implícita***Esta Função calcula a volatilidade implícita da opção***Function ImpVolCall(mktprice As Double, strike As Double, tenor As Double, _
spot As Double, fator As Double, err As Double)**tenor = tenor / 252**Volatility = 0.5**dv = err + 1**While Abs(dv) > err**d1 = Log(spot * fator / strike) + (0.5 * Volatility * Volatility) * tenor**d1 = d1 / (Volatility * Sqr(tenor))**d2 = d1 - Volatility * Sqr(tenor)**priceerror = spot * cdf(d1) - strike / fator * cdf(d2) - mktprice**vega = spot * Sqr(tenor / 3.1415926 / 2) * Exp(-0.5 * d1 * d1)**dv = priceerror / vega**Volatility = Volatility - dv**Wend**ImpVolCall = Volatility**End Function***Esta função aproxima a Distribuição Normal acumulada***Function cdf(x)**D = 1 / (1 + 0.33267 * Abs(x))**a = 0.4361836**b = -0.1201676**c = 0.937298**cdf = 1 - 1 / Sqr(2 * 3.1415926) * Exp(-0.5 * x * x) * _
(a * D + b * D * D + c * D * D * D)**If x < 0 Then cdf = 1 - cdf**End Function***Esta função calcula o delta de B-S.***Function BSDelta(spot As Double, strike As Double, vol As Double, _
tenor As Double, fator As Double)**tenor = tenor / 252**d1 = (Log(spot * fator / strike) + 0.5 * vol ^ 2 * tenor) / _
(vol * tenor ^ 0.5)**Nd1 = Application.NormSDist(d1)**BSDelta = Nd1**End Function*¹⁷ O código em Visual Basic utilizado foi adaptado de Wilmott(1998).

Apêndice 2: Hedge Dinâmico

No quadro abaixo, encontra-se os códigos em Visual Basic dos algoritmos utilizados para calcular os resultados do hedge dinâmico para os dois modelos.

<u>Esta função calcula o resultado do hedge dinâmico utilizando-se B-S</u>	
<i>Function BSDelta_Hedge(x, fim)</i>	
<i>i = x.Row</i>	'Posição inicial
<i>n = fim.Row</i>	'Posição Final onde o portfolio é liquidado
<i>If i = n Then bin = True Else bin = False</i>	
<i>cv = x.Value</i>	'Identifica se é compra ou venda de opção
<i>op = Cells(i, 3).Value</i>	'Preço de mercado da opção
<i>D = Cells(i, 21).Value</i>	'Delta de B-S
<i>spot = Cells(i, 4).Value</i>	'Preço da ação
<i>If cv = "V" Then</i>	'Calcula o Saldo Inicial da posição
<i>BSBalance = op - D * spot</i>	
<i>Else</i>	
<i>If cv = "C" Then</i>	
<i>BSBalance = -op + D * spot</i>	
<i>Else</i>	
<i>GoTo line1</i>	
<i>End If</i>	
<i>End If</i>	
<i>For i = i + 1 To n</i>	'Calcula o rebalanceamento do portfólio
<i>op = Cells(i, 3).Value</i>	
<i>d_ant = Cells(i - 1, 21).Value</i>	'Delta anterior de B-S
<i>d_atual = Cells(i, 21).Value</i>	'Delta atual de B-S
<i>spot = Cells(i, 4).Value</i>	
<i>cdi = Cells(i, 26).Value</i>	'Cdi utilizado (=0 se não for última operação do dia)
<i>If cv = "V" Then</i>	
<i>cashflow = -(d_atual - d_ant) * spot</i>	'Cálculo do fluxo de caixa das operações
<i>Else</i>	
<i>cashflow = (d_atual - d_ant) * spot</i>	
<i>End If</i>	
<i>BSBalance = BSBalance * (1 + cdi / 3000) + cashflow</i>	'Calcula o saldo das operações
<i>Next</i>	
<i>line1:</i>	
<i>If cv = "V" Then</i>	'Liquida as posições
<i>BSDelta_Hedge = BSBalance + d_atual * spot - op</i>	
<i>Else</i>	
<i>If cv = "C" Then</i>	
<i>BSDelta_Hedge = BSBalance - d_atual * spot + op</i>	

```

Else
  BSDelta_Hedge = 1000
End If
End If

If bin = True Then
  BSDelta_Hedge = 0
Else
  BSDelta_Hedge = BSDelta_Hedge
End If

End Function

```

Esta função calcula o resultado do hedge dinâmico utilizando-se B-S

Segue a mesma lógica da função anterior, só que utilizando os deltas implícitos na rede neural

```

Function NNDelta_Hedge(x, fim)
i = x.Row
n = fim.Row
If i = n Then bin = True Else bin = False
cv = x.Value
op = Cells(i, 3).Value
D = Cells(i, 20).Value
spot = Cells(i, 4).Value
If cv = "V" Then
  NNBalance = op - D * spot
Else
  If cv = "C" Then
    NNBalance = -op + D * spot
  Else
    GoTo line1
  End If
End If
For i = i + 1 To n
  op = Cells(i, 3).Value
  d_ant = Cells(i - 1, 20).Value
  d_atual = Cells(i, 20).Value
  spot = Cells(i, 4).Value
  cdi = Cells(i, 26).Value
  If cv = "V" Then
    cashflow = -(d_atual - d_ant) * spot
  Else
    cashflow = (d_atual - d_ant) * spot
  End If
  NNBalance = NNBalance * (1 + cdi / 3000) + cashflow

```

```
Next
line1:
If cv = "V" Then
  NNDelta_Hedge = NNBalance + d_atual * spot - op
Else
  If cv = "C" Then
    NNDelta_Hedge = NNBalance - d_atual * spot + op
  Else
    NNDelta_Hedge = 1000
  End If
End If

If bin = True Then
  NNDelta_Hedge = 0
Else
  NNDelta_Hedge = NNDelta_Hedge
End If

End Function
```

Referências Bibliográficas

- Abken, P. A., Nandi, S. “*Options and Volatility*”. Federal Reserve Bank of Atlanta. Economic Review. December 1996. 21-35.
- Apostolos-Paul “*Neural Networks in the Capital Markets*”. Wiley Finance Editions, 1995.
- Bakshi, Gurdip, Cao, Charles, Chen, Zhiwu “*Empirical Performance of Alternative Option Pricing Models*”. Journal of Finance, 52, 2003-2049.1997
- Becker, João Luiz., Lemgruber, Eduardo Facó, Felício, Rousely Freire “*Seguro Dinâmico de Portfólio*”. Revista Brasileira de Economia, 45(4), out/dez 1991, 629-647.
- Black, Fisher “*The Holes in Black and Scholes*”. Black-Scholes-Merton Supplement. Risk. September 1998. 6-10.
- Black, F., Scholes, M. “*The Pricing of Options and Corporate Liabilities*”, Journal of Political Economy, 81 (May-June 1973), 673-59
- Black, F., Scholes, M., “*The Valuation of Option Contratcts and a Test of Market Efficiency*”. Journal of Finance, 27, 399-418. 1972.
- Boudoukh, J., Richardson, M., Whitelaw, R. F. “Investigation of a Class of a Volatility Estimators”. The Journal of Derivatives. Spring 1997. 63-71.
- Brailsford, T. J., Faff, R. W. “An evaluation of Volatility Forecasting Techniques”. Journal of Banking and Finance. 1996. 419-438.

- Britten-Jones, Mark, Neuberger, Anthony “*Welcome to the real world*”. Black-Scholes-Merton Supplement. Risk. September 1998. 11-13.
- Brealey, Richard A., Myers, Stewart C. “*Principles of Corporate Finance*”. McGraw-Hill, Inc. 3rd edition, 1988.
- Campbell, John Y., Lo, Andrew W., MacKinlay, A.C. “*The Econometrics of Financial Markets*”. Princeton University Press, 1997.
- Chiras, D., Manaster, S. “*The Information Content of Option Prices and a Test of Market Efficiency*”. Journal of Financial Economics, 6, 213-34. September 1978.
- Cox, J. C., Ross, S. A., Rubinstein, M. “*Option Pricing: A Simplified Approach*”. Journal of Financial Economics, 1979, 229-263
- Cox, J. C., Ross, S. A.. “*The Valuation of Options for Alternatives Stochastic Processes*”, Journal of Financial Economics, 1976, 145-66.
- Cox, J. C., Rubinstein, M. “*A Survey of Alternative Option Pricing Models*”, Option Pricing. LexingtonBooks. Toronto. 1983
- Cox, J. C., Rubinstein, M. *Options Markets*. Prentice Hall. 1985
- Derman, E., Kamal, M. “*Correcting Black-Scholes*”. Risk, 82-85. January 1999.
- Dumas, Bernard, Fleming, Jeff, Whaley, Robert E. “*Implied Volatility Functions: Empirical Tests*”. Journal of Finance. Forthcoming issue, December 1998.
- Galai, D. “*Tests of Market Efficiency and The Chicago Board of Trade*”. Journal of Business, 50(April 1997), 167-97.

- Geske, R. “*The Valuation of Compound Options*”. *Journal of Financial Economics*, 1979. 63-81.
- Haykin, Simon. *Neural Networks: a comprehensive foundation*. Prentice Hall, Inc. 1999.
- Hull, J. C. *Options, Futures and Other Derivatives*. Upper Saddle River. Prentice Hall, Inc., 1997.
- Hull, J. White, A. “*Hull-White on Derivatives - A Compilation of Articles*”. Risk Publications. 1996.
- Hutchinson, J. M., Lo, A. W., Poggio, T. “*A NonParametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks*”. *Journal of Finance*. 1994. 851-889.
- Rubin, Rafael Levin “*Arbitragem de Volatilidade e a Essência da Equação de Black & Scholes*”. *Resenha BM&F -116*. 1998. 41-46.
- MacBeth, J.D., Merville, L.J. “*An Empirical Examination of the Black-Scholes Call Option Pricing Model*”. *Journal of Finance*, 34, 1173-86. December 1979.
- Merton, R. C. “*Option Pricing When Underlying Stock Returns are Discontinuous*”. *Journal of Financial Economics*, 1976. 125-44.
- Merton, R. C. “*Theory of Rational Option Pricing*”. *Bell Journal of Economics and Management Science*, 1973. 141-83.
- Neftci, Salih N.. *An Introduction to the Mathematics of Financial Derivatives*. Academic Press. 1996.

- Oksendal, B. *Astochastic Differential Equations*. Springer – Verlag. 1992.
- Refenes, Apostolos-Paul. *Neural Networks in the Capital Markets*. John Wiley & Sons, Inc. 1995.
- Rubinstein, M. “*Displaced Diffusion Option Pricing*”. *Journal of Finance*, 38, 213-17. March, 1983.
- Rubinstein, M. “*Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978*”. *Journal of Finance*, 40, 454-80. 1985.
- Rubinstein, M. “*Implied Binomial Tree*”. *Journal of Finance*, 49, 771-818. July, 1994.
- Smithson, Charles, Song, Shang “*Extended Family*”. *Black-Scholes-Merton Supplement. Risk*. September 1998. 14-17
- Stutzer, Michael. “*A Simple NonParametric Approach to Derivative Security Valuation*”. *Journal of Finance*. December 1996. 1633-1652.
- Taleb, Nassim. *Dynamic Hedging: Managing Vanilla and Exotic Options*. John Wiley & Sons, Inc. 1997.
- Trippi, Robert R., Turban, Efraim. *Neural Networks in Finance and Investing*. Probus Publishing Company. Chicago, 1993.
- Vasicek, O.A. “*An Equilibrium Characterization of the Term Structure Valuation*”. *Journal of Financial Economics*, 5. 1977. 177-188.
- Wilmott, Paul. *Derivatives: The theory and practice of financial engineering*. John Wiley & Sons, Inc. 1998.

