9
Referências

AIUBE, F. A. L. **Modelagem dos Preços Futuros de *Commodities*: Abordagem pelo Filtro de Partículas**. Tese de doutorado, Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2005.

AIUBE, F. A. L. **Processos Estocásticos em Finanças**. Disponível em (http://www.ind.puc-rio.br/UserFiles/File/aiube/Proc%20Estoc%20em%20Finanas.pdf).

ARNOLD, TOM; BERTUS, MARK J.; GODBEY, JONATHAN. **A Simplified Approach to Understanding The Kalman Filter Technique**. Engineering Economist, 2008, Vol. 53 Issue 2, p140-155.

BESSEMBINDER, H.; COUGHENOUR J. F.; SEGUIN P. J.; SMOLLER, M. M. **Mean reversion in equilibrium asset prices: evidence from the futures term structure**. The Journal of Finance, 1v, 1995.

BHATTACHARYA, S. **Project Valuation with Mean Reverting Cash Flow Streams**, Journal of. Finance 33(5) December, p. 1317-1331, 1978.

BLACK, F., SCHOLES, M. **The pricing of options and corporate liabilities**. Journal of Political Economy, 81, 637-654, 1973.

BOMHOFF, E. J. **Financial Forecasting for Business and Economics**. Academic Press Limited, San Diego, California, 1994.

BRANDÃO, L., (http://www.iag.puc-rio.br/~Brandão/Opcoes%20Reais/Dixit%20and%20Pindyck%20Files/Dixit%20and%20Pindyck.html).

BRENNAN, F. **The costs of convenience and the pricing of commodity contingent claims**. Unpublished paper, University of British Columbia, 1986.

BRENNAN, F.; SCHWARTZ, E.. **Evaluating natural resource investments**. Journal of Business 58, 135-157, 1985.

BRENNAN, M. J. **The price of convenience and the valuation of commodity contingent claims.** D. Lund, B. Oskendal, eds. Stochastic Models and Option Values: Application to Resources, Environment, and Investment Problems. North-Holland, New York, 1991.

CORTAZAR, G.; NARANJO, L. A multi-factor stochastic model for estimation procedure for the valuation and hedging of commodity contingent claims. Working paper, Pontificia Universidade Catolica de Chile, 2003.

CORTAZAR G.; SCHWARTZ, E. S.; NARANJO, L. Term structure estimation in low-frequency transaction markets: A Kalman filter approach with incomplete panel-data. Working paper, The Anderson School, University of California Los Angeles, 2003.

COX, J. C.; INGERSOLL, J. E.; ROSS, S. A. An Intertemporal General Equilibium Model of Asset Prices. Econometrica 53 363-384, 1985.

DIXIT, A. K., R. S. PINDYCK. **Investment Under Uncertainty**. Princeton University Press. Princeton, NJ, 1994.

DUFFIE, D. **Dynamic Asset Pricing Theory**. Princeton University Press, Pinceton, NJ, 1992.

FAMA, E.; FRENCH, K. **Business cycles and the behavior of metals prices.** Journal of Finance 43, 1075-1094, 1988.

FAMA, E.; FRENCH, K. **Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage**. Journal of Business 60, 55-74, 1987.

GIBSON, R., E. E. SCHWARTZ. ***Stochastic convinience yield and the pricing of oil contingent claims***. J. Finance 45, p. 959-976, 1990.

GIBSON, R.; SCHWARTZ, E. ***Valuation of long term oil-linked assets*.** Anderson Graduate School of Management, UCLA, Working Paper, #6-89, 1989.

GUJARATI, D. **Basic Econometrics**, 4th ed. McGraw-Hill Book Company, New York, 2003.

HARRISON, M.; KREPS, D. **Martingales and multiperiod securities markets**. Journal of Economic Theory, 1979.

HARRISON, M.; PLISKA S. **Martingales and stochastic integrals in theory of continuous trading**. Stochastic Processes and their Applications, 11, 313-316, 1981.

HARVEY, A. C. **Forecasting, structural time series models and the Kalman filter.** Cambridge University Press, Cambridge, 1989.

HOEL, P. G.; PORT, S. C.; STONE, C. J. **Introduction to probability theory**. Houghton Mifflin Company, Boston, 1971.

HULL, J. **Options, futures, and other derivatives**. 4th edition. Prentice Hall, 2000.

JOHNSTON, J.; DINARDO J. **Econometric methods**, 4th ed. McGraw-Hill Companies, New York, 1997.

LAUTIER, D. **The informational value of crude oil futures prices.** *Working paper*, Cereg, Université Paris IX, 2003.

MERTON, R. C. **Theory of rational option pricing**. Bell Journal of Economics and Management Science, 4, 141-183, 1973.

PADDOCK, J. L.; SIEGEL, D. R.; SMITH, J. L. **Option valuation of claims on real assets: The case of offshore petroleum leases**. Quarterly Journal of Economics 103, 479-509, 1988.

PILOPOVIC, D. **Energy Risk.** McGraw-Hill, New York, 1998.

ROSS, S. **Hedging long run commitments: Exercises in incomplete market pricing.** Banca Monte Econom. Notes 26 99-132, 1997.

SAMUELSON, P. A. **Proof that properly anticipated prices fluctuate randomly**. Industrial Management Review, 6, 41-49, Spring, 1965.

SCHWARTZ, E. S. Review of Investment Under Uncertainty. A.K. DIXIT, R. S. PINDYCK. Journal of Finance 50 1924-1928, 1995.

SCHWARTZ, E. S.; SMITH J. E. **Short term-variations and long-term dynamics in commodity prices.** Management Science, 46, 893-911. Addendum: Short-term variations and long-term dynamics in commodity prices: Incorporating a stochastic growth rate, 2000.

SHERIDAN, B. **600,000,000,000,000? It’s a number no one questions, but the size of derivatives market is not as shocking as it looks**. NEWSWEEK, Oct. 27, 2008.

SØRENSEN, C. Modeling seasonality in agricultural commodity futures. Journal of Futures Markets, 22, 393-426, 2002.

WEST, M.; HARRISON, J. **Bayesian Forecasting and Dynamic Models.** Second Edition. Springer, 1997.

10
Apêndice A – Aplicação do Lema de Itô ao Modelo de Gibson Schwartz (1990)

Através da teoria dos ativos contingentes baseada em argumentos de não arbitragem é possível se obter uma equação diferencial em termos do derivativo, que no caso é um contrato futuro.

Definindo o derivativo como B e sabendo que é função de S, δ e τ, onde $\tau = T - t$. Sabendo que B é duas vezes diferenciável em termos de S e δ, pode-se aplicar o Lema de Itô, e obter-se:

$$dB = \frac{\partial B}{\partial S} dS + \frac{\partial B}{\partial \delta} d\delta + \frac{\partial B}{\partial \tau} d\tau + \frac{1}{2}\frac{\partial^2 B}{\partial S^2}(dS)^2 + \frac{1}{2}\frac{\partial^2 B}{\partial \delta^2}(d\delta)^2 + \frac{\partial^2 B}{\partial S \partial \delta} dS d\delta$$

Como:

$$dS = \mu S dt + \sigma_1 S dz_1$$

Então:

$$(dS)^2 = \mu^2 S^2 dt^2 + 2\mu S \sigma_1 S dz_1 dt + {\sigma_1}^2 S^2 {dz_1}^2$$

Como dt é um período de tempo muito pequeno, dt^2 pode ser considerado zero. Sabe-se, ainda, que $dz_1 dt$ é igual à dt elevado a um número maior do que um e, portanto, pode-se considerar também igual a zero. É possível provar que dz^2 é igual a dt:

$$dz = \varepsilon_t \sqrt{dt}$$

$$E[dz] = \sqrt{dt} E[\varepsilon_t]$$

Onde,

$$\varepsilon_t \sim N(0,1) => E[\varepsilon_t] = 0 => E[dz] = 0$$

Então,

$$E[dz] = 0$$

Logo,

$$Var[dz] = E[(dz - E[dz])^2] = E[dz^2 - 2dzE[dz] + (E[dz])^2]$$

$$Var[dz] = E[dz^2]$$

Mas,

$$E[dz^2] = E\left[\left(\varepsilon_t \sqrt{dt}\right)^2\right] = dtE[(\varepsilon_t)^2]$$

E por definição,

$E[(\varepsilon_t)^2] = E[(\varepsilon_t - E[\varepsilon_t])^2] = Var[\varepsilon_t] = 1$

Portanto,

$E[dz^2] = dt$

Até aqui provou-se que o valor esperado de dz^2 é dt. O próximo passo é provar que a variância de dz^2 é zero. Se uma variável tem valor esperado igual a algum valor e variância zero pode-se afirmar que essa variável é igual ao próprio valor esperado. Aplicando a definição de variância.

$Var[dz^2] = E[(dz^2 - E[dz^2])^2] = E[dz^4 - 2dz^2E[dz^2] + E[dz^2]^2]$

Ou,

$Var[dz^2] = E[dz^4 - 2dz^2dt + dt^2] = E[dz^4] - 2dtE[dz^2] + dt^2$

Logo,

$Var[dz^2] = E[dz^4] - 2dt^2 + dt^2 = E[dz^4] - dt^2$

Onde,

$E[dz^4] = E[dz^2dz^2] = E[\varepsilon^2dt\varepsilon^2dt] = dt^2E[\varepsilon^4]$

No entanto, como a variável ε corresponde a uma normal padrão, é possível determinar-se o valor esperado acima da seguinte forma,

$E[\varepsilon^4] = 3$

Pois,

$E[\varepsilon^4] = E\left[\left(\varepsilon - E(\varepsilon)\right)^4\right] = E[\varepsilon^4]$

Ou seja, corresponde exatamente à definição do quarto momento da variável, ou ainda à curtose da distribuição. No entanto, sabe-se que o quarto momento da distribuição normal padrão é igual a três. Consequentemente, aplicando o resultado tem-se,

$E[dz^4] = 3dt^2$

Voltando,

$Var[dz^2] = 3dt^2 - dt^2 = 0$

Portanto, teremos:

$(dS)^2 = \mu^2S^2dt^2 + 2\mu S\sigma_1 Sdz_1dt + {\sigma_1}^2S^2{dz_1}^2 = {\sigma_1}^2S^2dt$

Elevando outro termo ao quadrado,

$(d\delta)^2 = (k(\alpha - \delta)dt + \sigma_2dz_2)(k(\alpha - \delta)dt + \sigma_2dz_2)$

Onde todos os termos são de ordem dt maior do que um, e por isso tendem a zero, exceto um. Logo, o resultado é

$(d\delta)^2 = \sigma_2^2 dt$

Para outro termo do Lema de Itô tem-se:

$dSd\delta = (\mu S dt + \sigma_1 S dz_1)(k(\alpha - \delta)dt + \sigma_2 dz_2)$

Neste caso todos os termos vão para zero, pois são da ordem de dt elevado a um número maior do que um, exceto um:

$dSd\delta = \sigma_1 \sigma_2 S \rho dt$

Como:

$\tau = T - t => d\tau = -dt$

Substituindo na equação do Lema de Itô as relações acima e considerando as seguintes abreviações para as derivadas parciais teremos:

$$dB = B_S(\mu S dt + \sigma_1 S dz_1) + B_\delta(\mathrm{k}(\alpha - \delta)\mathrm{dt} + \sigma_2 \mathrm{dz}_2) + \mathrm{B}_\tau(-\mathrm{dt}) + \frac{1}{2}\mathrm{B}_{SS}({\sigma_1}^2 S^2 dt) + \frac{1}{2} B_{\delta\delta}\sigma_2^2 dt + B_{S\delta}\sigma_1 \sigma_2 S \rho dt$$

Onde:

$$B_S = \frac{\partial B}{\partial S}$$

$$B_\delta = \frac{\partial B}{\partial \delta}$$

$$B_\tau = \frac{\partial B}{\partial \tau}$$

$$B_{SS} = \frac{\partial^2 B}{\partial S^2}$$

$$B_{\delta\delta} = \frac{\partial^2 B}{\partial \delta^2}$$

$$B_{S\delta} = \frac{\partial^2 B}{\partial S \delta}$$

Agrupando os termos em dt:

$$dB = \left[-\mathrm{B}_\tau + \frac{1}{2}\mathrm{B}_{SS}{\sigma_1}^2 S^2 + B_{S\delta} S \rho \sigma_1 \sigma_2 + \frac{1}{2} B_{\delta\delta}\sigma_2^2 + B_S \mu S + B_\delta \mathrm{k}(\alpha - \delta)\right] dt + \sigma_1 S B_S dz_1 + \sigma_2 \mathrm{B}_\delta \mathrm{dz}_2$$

Considerando a taxa de juros sem risco como determinística e condições de mercado perfeito, a equação diferencial obtida pelo método dos ativos contingentes, que considera condições de não arbitragem, será:

$$\frac{1}{2} B_{SS} S^2 {\sigma_1}^2 + \frac{1}{2} B_{\delta\delta}\sigma_2^2 + B_{S\delta} S \rho \sigma_1 \sigma_2 + B_S S(r - \delta) + B_\delta[k(\alpha - \delta) - \lambda\sigma_2] - B_\tau - rB = 0$$

Sujeito à condição inicial de que (para um contrato futuro):

$F(S, \delta, 0) = S$

Onde:

λ= preço por unidade de risco da taxa de retorno de conveniência

F representa um contrato futuro.

11
Apêndice B – Função Geradora de Momentos

A função geradora de momentos é dada pelo valor esperado de e^{tx}:

$M_X(t) = E[e^{tX}]$

Portanto,

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

Para o caso em que a distribuição de probabilidade da variável X é uma Normal, dada por:

$X \sim N(\mu, \sigma^2)$

Ou seja,

$$f(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Então a função geradora de momentos será

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

Ou,

$$M_X(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{\frac{2\sigma^2 tx}{2\sigma^2}} e^{\frac{-(x^2-2x\mu+\mu^2)}{2\sigma^2}} dx$$

Desenvolvendo,

$$M_X(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{\frac{-(x^2-2(\mu+t\sigma^2)x+\mu^2)}{2\sigma^2}} dx$$

Se completarmos o quadrado no expoente,

$$(x^2 - 2(\mu + t\sigma^2)x + \mu^2) = [x - (\mu + t\sigma^2)]^2 - 2\mu t\sigma^2 - t^2\sigma^4$$

Substituindo,

$$M_X(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{\frac{-\left\{[x-(\mu+t\sigma^2)]^2-2\mu t\sigma^2-t^2\sigma^4\right\}}{2\sigma^2}} dx$$

Ou,

$$M_X(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{\mu t+\frac{1}{2}t^2\sigma^2} \int_{-\infty}^{\infty} e^{\frac{-\left\{[x-(\mu+t\sigma^2)]^2\right\}}{2\sigma^2}} dx$$

Chamando de $[x - (\mu + t\sigma^2)]/\sigma$ de u, então $dx = \sigma du$. Portanto

$$M_X(t) = \frac{1}{\sqrt{2\pi}} e^{\mu t + \frac{1}{2}t^2\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2}\, du$$

Se o termo $\frac{1}{\sqrt{2\pi}}$ for colocado novamente dentro da integral, observar-se-á que a integral corresponde exatamente à toda a área abaixo da função de densidade da Normal Padrão, e portanto é exatamente igual a um. Logo,

$$M_X(t) = e^{\mu t + \frac{1}{2}t^2\sigma^2}$$

O r-ésimo momento em torno da origem será dado por:

$$\mu_r' = E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx$$

Essa definição permite duas observações. A primeira em relação à média, pois sabe-se que

$$E(X) = \mu$$

O que quer dizer que o primeiro momento, quando substitui-se r por um, é exatamente a definição de média, que chamou-se de μ.

A segunda constatação permite que seja feita uma relação entre a variância e o segundo momento. Substituindo r na definição acima por dois, tem-se o segundo momento

$$\mu_2' = E(X^2)$$

No entanto, uma das formas de definir-se a variância é

$$\sigma^2 = E(X^2) - [E(X)]^2$$

Portanto, a variância poderá ser também definida como uma função do segundo momento

$$\sigma^2 = \mu_2' - \mu^2$$

Outra definição bastante útil diz que se X tiver função geradora de momentos $M_X(t)$, então

$$\mu_r' = \left.\frac{d^r M_X(t)}{dt^r}\right|_{t=0}$$

Aplicando a definição acima à função encontrada para uma variável com distribuição Normal

$$\left.\frac{dM_X(t)}{dt}\right|_{t=0} = \left. e^{\mu t + \frac{1}{2}t^2\sigma^2}(\mu + t\sigma^2)\right|_{t=0}$$

Logo,

$$\left.\frac{dM_X(t)}{dt}\right|_{t=0} = \mu$$

E o segundo momento será dado por

$$\left.\frac{d^2M_X(t)}{dt^2}\right|_{t=0} = \left.\frac{d\left[e^{\mu t+\frac{1}{2}t^2\sigma^2}(\mu+t\sigma^2)\right]}{dt}\right|_{t=0}$$

$$\left.\frac{d^2M_X(t)}{dt^2}\right|_{t=0} = \left.e^{\mu t+\frac{1}{2}t^2\sigma^2}(\mu+t\sigma^2)(\mu+t\sigma^2)+e^{\mu t+\frac{1}{2}t^2\sigma^2}(\sigma^2)\right|_{t=0}$$

$$\left.\frac{d^2M_X(t)}{dt^2}\right|_{t=0} = \mu^2+\sigma^2$$

Com os resultados obtidos acima é possível determinar a média e a variância de uma distribuição Normal. A média será igual ao primeiro momento. Isolando a variância no resultado acima tem-se

$$\sigma^2 = \mu_2' - \mu^2$$

No entanto, determinou-se que a variância se relaciona com o segundo momento da seguinte forma

$$\mu_2' = \sigma^2 + \mu^2$$

Substituindo a equação acima na anterior

$$\sigma^2 = \sigma^2 + \mu^2 - \mu^2$$

Consequentemente, a variância da normal é igual ao termo σ da sua função de distribuição de probabilidade elevado ao quadrado.

11.1. Aplicação da Função Geradora de Momentos ao Modelo de S&S (2000)

No caso do modelo de S&S (2000) a distribuição de probabilidade da função que se deseja aplicar a função geradora de momentos é Normal. Por isso, poder-se-á utilizar a fórmula encontrada acima, tal que

$$M_X(t) = E(e^{tX}) = e^{\mu t+\frac{1}{2}t^2\sigma^2}$$

No caso aqui proposto

$X = ln(S_t)$ e $t = 1$

A média e a variância da distribuição normal são dadas por μ e σ^2 respectivamente. Logo,

$$E\left[e^{ln(S_t)}\right] = exp\left(E[ln(S_t)] + \frac{1}{2}Var[ln(S_t)]\right)$$

$$E[S_t] = exp\left(E[ln(S_t)] + \frac{1}{2}Var[ln(S_t)]\right)$$

$$ln\ (E[S_t]) = E[ln(S_t)] + \frac{1}{2}Var[ln(S_t)]$$

Desta forma, encontrou-se uma relação que fornece o logaritmo do valor esperado do preço à vista em função do valor esperado e da variância do logaritmo do preço à vista. Essas variáveis já foram encontradas para o modelo e por isso é possível obter-se o logaritmo do valor esperado do preço à vista.

11.2. Aplicação da Função Geradora de Momentos ao Modelo de S&S (2000) Sob Processos Neutros ao Risco

A demonstração é bastante parecida com a anterior em que as medidas de probabilidade não eram as do mundo sem risco. Como $ln(S_t)$ continua sendo igual à soma das variáveis de estado, que continuam tendo distribuições normais, apenas com as médias deslocadas, $ln(S_t)$ também continua sendo Normalmente distribuído. Desta forma, ao aplicar-se a função geradora de momentos, obtêm-se o mesmo resultado

$$M_X(t) = E^*(e^{tX}) = e^{\mu t + \frac{1}{2}t^2\sigma^2}$$

Substituindo na função geradora de momentos as relações para μ e σ^2, sabendo que $X = ln(S_t)$ e aplicando o valor de $t = 1$

$$E^*\left[e^{ln(S_t)}\right] = exp\left(E^*[ln(S_t)] + \frac{1}{2}Var^*[ln(S_t)]\right)$$

$$E^*[S_t] = exp\left(E^*[ln(S_t)] + \frac{1}{2}Var^*[ln(S_t)]\right)$$

$$ln\,(E^*[S_t]) = E^*[ln(S_t)] + \frac{1}{2}Var^*[ln(S_t)]$$

12
Apêndice C – Demonstrações das Propriedades Estatísticas dos Processos Estocásticos Abordados

As próximas seções demonstram como obter o valor esperado e a variância dos processos estocásticos abordados nesta dissertação, ou seja, o processo de Wiener, o movimento geométrico browniano e o movimento de reversão à média.

12.1. Demonstração da Evolução da Variância do Processo de Wiener no Tempo

A demonstração de que a variância do processo de Wiener é crescente no tempo pode ser feita da seguinte forma:

$Var(a) = E[(a - E[a])^2]$

Portanto, para o processo de Wiener:

$Var[z(T) - z(0)] = E[(z(T) - z(0) - E[z(T) - z(0)])^2]$

Onde:

$z(T)$ = ao estado do processo no tempo T

$z(0)$ = ao estado do processo no tempo 0

Podemos ver $z(T) - z(0)$ como um somatório de variações dentro do intervalo entre zero e T. Desta forma, as variações seriam:

$z(1) - z(0) = \varepsilon_1\sqrt{\Delta t}$

$z(2) - z(1) = \varepsilon_2\sqrt{\Delta t}$

$\vdots$

$z(T) - z(T-1) = \varepsilon_N\sqrt{\Delta t}$

Onde:

$N = \dfrac{T}{\Delta t}$

$t = 0, 1, 2, \dots, T$

O somatório das variações será:

$z(T) - z(0) = \varepsilon_1\sqrt{\Delta t} + \varepsilon_2\sqrt{\Delta t} + \cdots + \varepsilon_N\sqrt{\Delta t} = \sum_{i=1}^{N} \varepsilon_i\sqrt{\Delta t}$

Logo,

$$E[z(T)-z(0)] = E\left[\sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}\right] = \sum_{i=1}^{N} E[\varepsilon_i]\sqrt{\Delta t} = 0$$

Então,

$$Var[z(T)-z(0)] = E[(z(T)-z(0)-0)^2] = E\left[\left(\sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}\right)^2\right]$$

Mas,

$$\left(\sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}\right)^2 = (\varepsilon_1\sqrt{\Delta t} + \varepsilon_2\sqrt{\Delta t} + \cdots + \varepsilon_N\sqrt{\Delta t})(\varepsilon_1\sqrt{\Delta t} + \varepsilon_2\sqrt{\Delta t} + \cdots + +\varepsilon_N\sqrt{\Delta t})$$

Desenvolvendo,

$$\left(\sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}\right)^2 = \varepsilon_1^2\Delta t + \varepsilon_1\varepsilon_2\,\Delta t + \cdots + \varepsilon_1\varepsilon_N\Delta t + \varepsilon_2\varepsilon_1\Delta t + \varepsilon_2^2\Delta t + \cdots + \varepsilon_2\varepsilon_N\Delta t + \cdots + \varepsilon_N\varepsilon_1\,\Delta t + \varepsilon_N\varepsilon_2\,\Delta t + \cdots + \varepsilon_N^2\Delta t$$

Tirando o valor esperado do somatório acima:

$$E\left[\left(\sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}\right)^2\right] = E[\varepsilon_1^2\Delta t + \varepsilon_1\varepsilon_2\,\Delta t + \cdots + \varepsilon_1\varepsilon_N\Delta t + \varepsilon_2\varepsilon_1\Delta t + +\varepsilon_2^2\Delta t + \cdots + \varepsilon_2\varepsilon_N\Delta t + \cdots + \varepsilon_N\varepsilon_1\,\Delta t + \varepsilon_N\varepsilon_2\,\Delta t + \cdots + \varepsilon_N^2\Delta t]$$

Rearranjando

$$E\left[\left(\sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}\right)^2\right] = \Delta t(E[\varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_N^2 + \varepsilon_1\varepsilon_2 + \cdots + \varepsilon_N\varepsilon_{N-1}])$$

Como o valor esperado da soma é igual à soma dos valores esperados

$$E\left[\left(\sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}\right)^2\right] = \Delta t(E[\varepsilon_1^2] + E[\varepsilon_2^2] + \cdots + E[\varepsilon_N^2] + E[\varepsilon_1\varepsilon_2] + \cdots + E[\varepsilon_N\varepsilon_{N-1}])$$

Todos os termos de $E[\varepsilon_i\varepsilon_j]$ $onde\ i \neq j$ serão zero.

$$E\left[\left(\sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}\right)^2\right] = \Delta t(E[\varepsilon_1^2] + E[\varepsilon_2^2] + \cdots + E[\varepsilon_N^2])$$

Onde:

$$E[\varepsilon_i^2] = Var[\varepsilon_i] = E[(\varepsilon_i - E[\varepsilon_i])^2] = E[\varepsilon_i^2]$$

Então,

$$E\left[\left(\sum_{i=1}^{N} \varepsilon_i \sqrt{\Delta t}\right)^2\right] = \Delta t(1 + 1 + \cdots + 1) = \Delta tN$$

Portanto,

$$Var[z(T)-z(0)] = \Delta tN$$

Como:

$$N = \frac{T}{\Delta t}$$

Então,

$$Var[z(T) - z(0)] = \Delta t \frac{T}{\Delta t} = T$$

Desta forma, conclui-se que a variação esperada entre zero e T é zero e a variância é T. Podemos dizer, portanto, que:

$$z(T) - z(0) \sim N(0, T)$$

12.2. Demonstração do Valor Esperado e da Variância do MGB

A média e a variância do incremento do MGB serão calculadas através do Lema de Itô. O Lema de Itô pode ser comparado a uma Série de Taylor na qual os termos em dt, de ordem maior do que um, podem ser desprezados. Assumindo que $x(t)$ segue uma distribuição LogNormal, o que é extremamente coerente para modelagem de preços de ativos, pois não podem assumir valores negativos, criaremos uma nova variável tal que:

$$F(x) = ln\,(x)$$

Pelo Lema de Itô:

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial t} dt + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(dx)^2 + \frac{1}{6}\frac{\partial^2 F}{\partial x^3}(dx)^3 + \cdots$$

Onde,

$$(dx)^2 = \big(a(x,t)\big)^2 (dt)^2 + 2a(x,t)b(x,t)dtdz + \big(b(x,t)\big)^2 (dz)^2$$

Onde,

$$dtdz = (dt)^{3/2}$$

$$(dz)^2 = dt$$

Logo,

$$(dx)^2 = \big(b(x,t)\big)^2 dt$$

Todos os termos de $(dx)^3$ vão ser de ordem dt maior que um, portanto vão para zero.

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial t} dt + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}(dx)^2$$

Onde:

$$\frac{\partial F}{\partial x} = \frac{1}{x}$$

$$\frac{\partial F}{\partial t} = 0$$

$$\frac{\partial^2 F}{\partial x^2} = -\frac{1}{x^2}$$

$dx = \alpha x dt + \sigma x dz$

$(dx)^2 = \left(b(x,t)\right)^2 dt = \sigma^2 x^2 dt$

Então,

$dF = \frac{1}{x}(\alpha x dt + \sigma x dz) - \frac{1}{2}\frac{1}{x^2}\sigma^2 x^2 dt$

Colocando dt em evidência

$dF = \left(\alpha - \frac{1}{2}\sigma^2\right)dt + \sigma dz$

Logo,

$dF \sim N\left[\left(\alpha - \frac{1}{2}\sigma^2\right)dt, \sigma^2 dt\right]$

Considerando um intervalo de tempo $(0, T)$:

$F(x_t) - F(x_0) \sim N\left[\left(\alpha - \frac{1}{2}\sigma^2\right)T, \sigma^2 T\right]$

Ou,

$F(x_t) \sim N\left[F(x_0) + \left(\alpha - \frac{1}{2}\sigma^2\right)T, \sigma^2 T\right]$

Como $F(x_t) = ln(x_t)$, então,

$ln(x_t) \sim N\left[F(x_0) + \left(\alpha - \frac{1}{2}\sigma^2\right)T, \sigma^2 T\right] \sim N[u, v^2]$

Considerando,

$Y = ln(x_t) \sim N[u, v^2]$

Y terá uma função distribuição de probabilidade igual a:

$f(Y) = \frac{1}{\sqrt{2\pi v^2}} e^{-\frac{1}{2}\left(\frac{Y-u}{v}\right)^2}$

Como,

$Y = ln(x_t) => x_t = e^Y$

Mais uma vez precisaremos do conceito de função geradora de momentos:

$M_Y(k) = E[e^{kY}] = \int_{-\infty}^{+\infty} e^{kY} f(Y) dY$

Substituindo o valor de $f(Y)$ dado pela função de distribuição de probabilidade na função geradora de momento:

$E[e^{kY}] = \int_{-\infty}^{+\infty} e^{kY}\left[\frac{1}{\sqrt{2\pi v^2}} e^{-\frac{1}{2}\left(\frac{Y-u}{v}\right)^2}\right] dY = \int_{-\infty}^{+\infty}\left[\frac{1}{\sqrt{2\pi v^2}} e^{kY-\frac{1}{2}\left(\frac{Y-u}{v}\right)^2}\right] dY$

$E[e^{kY}] = \int_{-\infty}^{+\infty}\left[\frac{1}{\sqrt{2\pi v^2}} e^{\frac{2v^2 kY - (Y-u)^2}{2v^2}}\right] dY$

Completando o quadrado do numerador do expoente para obtermos algo do tipo $e^Y e^W$:

$$2v^2kY-(Y-u)^2=2v^2kY-(Y-u)^2+(v^4k^2-v^4k^2)+(2v^2ku-2v^2ku)$$

$$2v^2kY-(Y-u)^2=-v^4k^2+2v^2kY-2v^2ku-(Y-u)^2+v^4k^2+2v^2ku$$

$$2v^2kY-(Y-u)^2=-[v^4k^2-2v^2k(Y-u)+(Y-u)^2]+v^4k^2+2v^2ku$$

$$2v^2kY-(Y-u)^2=-[v^2k-(Y-u)]^2+v^4k^2+2v^2ku$$

$$2v^2kY-(Y-u)^2=-[Y-(u+v^2k)]^2+v^4k^2+2v^2ku$$

Portanto,

$$E[e^{kY}]=\int_{-\infty}^{+\infty}\left[\frac{1}{\sqrt{2\pi v^2}}e^{\frac{-[Y-(u+v^2k)]^2}{2v^2}}e^{\frac{v^4k^2+2v^2ku}{2v^2}}\right]dY$$

$$E[e^{kY}]=e^{\frac{v^4k^2+2v^2ku}{2v^2}}\int_{-\infty}^{+\infty}\left[\frac{1}{\sqrt{2\pi v^2}}e^{\frac{-[Y-(u+v^2k)]^2}{2v^2}}\right]dY$$

$$E[e^{kY}]=e^{\frac{v^2k^2}{2}+ku}\int_{-\infty}^{+\infty}\left[\frac{1}{\sqrt{2\pi v^2}}e^{\frac{-[Y-(u+v^2k)]^2}{2v^2}}\right]dY$$

Realizando uma troca de variáveis onde:

$$w=\frac{Y-(u+v^2k)}{v}$$

Então, derivando w em função de Y

$$\frac{dw}{dY}=\frac{1}{v}=>dY=vdw$$

Voltando,

$$E[e^{kY}]=e^{\frac{v^2k^2}{2}+ku}\int_{-\infty}^{+\infty}\left[\frac{1}{\sqrt{2\pi v^2}}e^{\frac{-[w]^2}{2}}\right]vdw=e^{\frac{v^2k^2}{2}+ku}$$

A integral é equivalente à função densidade de probabilidade de uma Normal padrão, com média zero e variância um. Desta forma, a integral é igual a um, pois corresponde à probabilidade de ocorrência de todos os valores possíveis, ou seja, corresponde a cem por cento ou um.

Substituindo os valores de Y, v e u e considerando k igual a um por ser o cálculo do primeiro momento que corresponde à média:

$$E\left[e^{ln(x_t)}\right]=e^{\frac{\sigma^2T}{2}+F(x_0)+\left(\alpha-\frac{1}{2}\sigma^2\right)T}=e^{\frac{\sigma^2T}{2}+ln(x_0)+\left(\alpha-\frac{1}{2}\sigma^2\right)T}$$

Ou,

$$E\left[e^{ln(x_t)}\right]=x_0e^{\frac{\sigma^2T}{2}+\left(\alpha-\frac{1}{2}\sigma^2\right)T}$$

Então,

$E[x_t] = x_0 \mathrm{e}^{\alpha T}$

A variância de x_t pode ser definida como:

$Var[x_t] = E[(x_t - E[x_t])^2] = E[{x_t}^2 - 2x_t E[x_t] + (E[x_t])^2]$

Ou,

$Var[x_t] = E[{x_t}^2] - 2(E[x_t])^2 + (E[x_t])^2 = E[{x_t}^2] - (E[x_t])^2$

Onde:

$(E[x_t])^2 = (x_0 e^{\alpha T})^2 = {x_0}^2 e^{2\alpha T}$

E,

$E[{x_t}^2] = E\left[e^{(ln(x_t))^2}\right] = E\left[e^{2ln(x_t)}\right] = E[e^{2Y}] = e^{\frac{v^2 2^2}{2}+2u} = e^{2(v^2+u)}$

Substituindo,

$E[{x_t}^2] = e^{2\left(\sigma^2 T + F(x_0) + \left(\alpha - \frac{1}{2}\sigma^2\right)T\right)} = e^{2\left(ln(x_0) + \alpha T + \frac{1}{2}\sigma^2 T\right)} = {x_0}^2 e^{2\alpha T + \sigma^2 T}$

Então,

$Var[x_t] = {x_0}^2 \mathrm{e}^{2\alpha T + \sigma^2 T} - {x_0}^2 \mathrm{e}^{2\alpha T} = {x_0}^2 \mathrm{e}^{2\alpha T}\left(\mathrm{e}^{\sigma^2 T} - 1\right)$

12.3. Média, Variância e Outras Características do MRM

Seja o MRM dado por

$dx = \eta(\bar{x} - x)dt + \sigma dz$

Onde:

η = velocidade de reversão à média

$\bar{x}$ = média de longo prazo

Para calcularmos o valor esperado e a variância de x_t iremos utilizar um fator de integração $e^{\eta t}$e multiplicar todos os termos da equação acima por ele:

$e^{\eta t} dx = e^{\eta t}\eta(\bar{x} - x)dt + e^{\eta t}\sigma dz$

Definindo uma variável,

$g(t, x) = e^{\eta t} x$

Aplicando o Lema de Itô a $g(t, x)$:

$dg(t, x) = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial t} dt + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dx)^2$

Onde:

$\frac{\partial g}{\partial x} = e^{\eta t}$

$$\frac{\partial g}{\partial t} = e^{\eta t} x\, \eta$$

$$\frac{\partial^2 g}{\partial x^2} = 0$$

Substituindo em na equação obtida com o Lema de Itô as relações acima:

$$d(e^{\eta t} x) = e^{\eta t}(\eta(\bar{x} - x)dt + \sigma dz) + e^{\eta t} x\, \eta dt$$

Abrindo o termo entre parênteses,

$$d(e^{\eta t} x) = \eta e^{\eta t} \bar{x} dt - \eta e^{\eta t} x dt + \sigma e^{\eta t} dz + e^{\eta t} x\, \eta dt$$

Simplificando,

$$d(e^{\eta t} x) = \eta e^{\eta t} \bar{x} dt + \sigma e^{\eta t} dz$$

Integrando os termos de zero a t:

$$\int_{s=0}^{t} d(e^{\eta s} x_s) ds = \int_{s=0}^{t} \eta e^{\eta s} \bar{x} ds + \int_{s=0}^{t} \sigma e^{\eta s} dz_s$$

$$e^{\eta s} x_s \Big|_{s=0}^{t} = \bar{x} e^{\eta s} \Big|_{s=0}^{t} + \sigma \int_{s=0}^{t} e^{\eta s} dz_s$$

$$e^{\eta t} x_t - x_0 = \bar{x}(e^{\eta t} - 1) + \sigma \int_{s=0}^{t} e^{\eta s} dz_s$$

Dividindo a equação por $e^{\eta t}$,

$$x_t = x_0 e^{-\eta t} + \bar{x}(1 - e^{-\eta t}) + \sigma \int_{s=0}^{t} e^{\eta(s-t)} dz_s$$

O valor esperado e a variância do MRM serão:

$$E[x_t] = x_0 e^{-\eta t} + \bar{x}(1 - e^{-\eta t}) + \sigma \int_{s=0}^{t} e^{\eta(s-t)} E[dz_s]$$

A única variável aleatória do lado direito da equação é o incremento de Wiener, e como ele possui valor esperado igual a zero:

$$E[x_t] = x_0 e^{-\eta t} + \bar{x}(1 - e^{-\eta t})$$

Ou,

$$E[x_t] = \bar{x} + e^{-\eta t}(x_0 - \bar{x})$$

No artigo de SCHWARTZ E SMITH (2000) o processo de reversão à média tende para zero. Para chegarmos ao valor de média citado no artigo para o MRM, basta considerarmos na equação acima que $\bar{x} = 0$, assim teremos:

$$E[x_t] = x_0 e^{-\eta t}$$

Quando os autores consideram o MRM neutro ao risco o processo passa a ser dado pela equação:

$$d\chi_t = \left(-k\chi_t - \lambda_\chi\right)dt + \sigma_\chi dz_\chi^*$$

Em que a tendência do processo ao invés de zero passa a ser $-\frac{\lambda_\chi}{k}$. Neste caso o valor esperado do processo será:

$E[x_t] = x_0 e^{-kt} - \frac{\lambda_\chi}{k}(1 - e^{-kt})$

A variância do MRM pode ser calculada da seguinte maneira:

$$Var[x_t] = E[(x_0 e^{-\eta t} + \bar{x}(1 - e^{-\eta t}) + \sigma \int_{s=0}^{t} e^{\eta(s-t)} dz_s - x_0 e^{-\eta t} - \bar{x}(1 - + e^{-\eta t}))^2]$$

Simplificando,

$$Var[x_t] = E\left[\left(\sigma \int_{s=0}^{t} e^{\eta(s-t)} dz_s\right)^2\right]$$

Elevando ao quadrado e retirando do valor esperado o termo constante,

$$Var[x_t] = \sigma^2 E\left[\int_{s=0}^{t} \left(e^{\eta(s-t)}\right)^2 (dz_s)^2\right]$$

Como o incremento de Wiener ao quadrado é igual a dt,

$$Var[x_t] = \sigma^2 \int_{s=0}^{t} e^{2\eta(s-t)} ds$$

Resolvendo a integral,

$$Var[x_t] = \sigma^2 \left.\frac{e^{2\eta(s-t)}}{2\eta}\right|_{s=0}^{t}$$

Substituindo os limites no resultado encontrado,

$$Var[x_t] = \frac{\sigma^2}{2\eta}(1 - e^{-2\eta t})$$

É interessante observar-se algumas características do MRM quando varia-se o tempo e a velocidade de reversão à média.

$$lim_{t\to\infty} E[x_t] = lim_{t\to\infty} \bar{x} + e^{-\eta t}(x_0 - \bar{x})$$

Quando t vai para infinito, $e^{-\eta t}$ vai para zero, o que acarreta com que $e^{-\eta t}(x_0 - \bar{x})$ também vá para zero. Portanto,

$$lim_{t\to\infty} E[x_t] = \bar{x}$$

Quanto à variância do processo,

$$lim_{t\to\infty} Var[x_t] = \lim_{t\to\infty} \frac{\sigma^2}{2\eta}(1 - e^{-2\eta t})$$

Como $e^{-2\eta t}$ vai para zero quando t tende a infinito, tem-se que $(1 - e^{-2\eta t})$ vai para um. Portanto,

$$lim_{t\to\infty} Var[x_t] = \frac{\sigma^2}{2\eta}$$

Observando, agora, variações na velocidade de reversão à média, pode-se observar que quando $\eta \to \infty$, tem-se

$$lim_{\eta\to\infty} E[x_t] = lim_{\eta\to\infty} \bar{x} + e^{-\eta t}(x_0 - \bar{x})$$

Assim, $\frac{1}{e^{\eta t}}(x_0 - \bar{x})$ tenderá a zero, pois, $e^{\eta t}$ quando η vai para infinito tende a infinito, e portanto, $\frac{1}{e^{\eta t}}$ tende a zero. Então,

$$\lim_{\eta\to\infty} E[x_t] = \bar{x}$$

A variância do processo, quando $\eta \to \infty$, se comporta da seguinte maneira:

$$lim_{\eta\to\infty} Var[x_t] = lim_{\eta\to\infty} \frac{\sigma^2}{2\eta}(1 - e^{-2\eta t})$$

O termo $(1 - e^{-2\eta t})$ vai para um, e o outro, $\frac{\sigma^2}{2\eta}$, tende a zero. Portanto, a variância do processo será:

$$lim_{\eta\to\infty} Var[x_t] = 0$$

As equações acima demonstram que quando a velocidade de reversão à média é suficientemente grande, o processo não se desvia de sua média de longo prazo.

Ao tender-se a velocidade de reversão à média a zero, tem-se que o valor esperado será dado por:

$$lim_{\eta\to 0} E[x_t] = lim_{\eta\to 0}\, \bar{x} + e^{-\eta t}(x_0 - \bar{x})$$

Como $e^{-\eta t}$ vai para um, tem-se:

$$lim_{\eta\to 0} E[x_t] = x_0$$

E a variância do processo será dada por:

$$lim_{\eta\to 0} Var[x_t] = lim_{\eta\to 0} \frac{\sigma^2}{2\eta}(1 - e^{-2\eta t})$$

Sabe-se que:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

Quando x vai para zero x^i vai mais rapidamente se $i = 1, 2, ...$, e portanto pode-se desconsiderar alguns termos, ficando com a igualdade:

$$e^x = 1 + x$$

Substituindo o resultado obtido acima na equação anterior, e considerando que $x = -2\eta t$, tem-se:

$$lim_{\eta\to 0} Var[x_t] = lim_{\eta\to 0} \frac{\sigma^2}{2\eta}\big(1 - (1 - 2\eta t)\big)$$

$$lim_{\eta\to 0} Var[x_t] = lim_{\eta\to 0} \frac{\sigma^2}{2\eta}(2\eta t)$$

$$\underset{\eta\to 0}{lim} Var[x_t] = \sigma^2 t$$

Comparando a variância e o valor esperado do MRM, quando $\eta \to 0$, com a de um Movimento Browniano sem *drift* constata-se que eles são exatamente iguais. Outra forma de se ver essa característica do MRM é igualar-se na sua equação o valor de η a zero.

13
Apêndice D – Minimização do Erro Quadrático

Para que a minimização do erro quadrático seja realizada será necessário que sejam atendidas duas condições de otimização. A condição de primeira ordem garante que o ponto encontrado seja crítico. A condição de segunda ordem é aquela que garante que o ponto além de crítico seja de mínimo.

13.1. Condição de 1ª Ordem

O método de estimação por mínimos quadrados ordinários será realizado através da minimização do erro quadrático. Para isso é necessário que conheça-se o erro quadrático, o qual será encontrado a seguir.

$$\boldsymbol{e_t = Y_t - X_t b_t}$$

A partir deste momento o índice t será suprimido por questões de simplificação. Logo, o vetor de erros transposto será

$$\boldsymbol{e' = Y' - b'X'}$$

E o quadrado dos erros

$$\boldsymbol{e'e} = (\boldsymbol{Y' - b'X'})(\boldsymbol{Y - Xb})$$

Desenvolvendo

$$\boldsymbol{e'e = Y'Y - Y'Xb - b'X'Y + b'X'Xb}$$

Como,

$$\boldsymbol{Y'Xb} = (\boldsymbol{Y'Xb})'$$

Pois trata-se de um escalar, e,

$$(\boldsymbol{Y'Xb})' = \boldsymbol{b'X'Y}$$

Então,

$$\boldsymbol{e'e = Y'Y - 2Y'Xb + b'X'Xb}$$

A condição de 1ª ordem para a minimização do erro quadrático é:

$$\boldsymbol{\frac{\partial e'e}{\partial b} = 0}$$

Aplicando a condição anterior na equação acima do erro quadrático, e derivando cada termo:

$$\frac{\partial \boldsymbol{Y'Y}}{\partial \boldsymbol{b}} = \boldsymbol{0}$$

$$\frac{\partial \boldsymbol{Y'Xb}}{\partial \boldsymbol{b}} = \cdots$$

$$\boldsymbol{Y'X} = [Y_1 \quad Y_2 \quad \dots \quad Y_T]\begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{kn} \end{bmatrix}$$

$$\boldsymbol{Y'X} = [Y_1 + Y_2 + \cdots + Y_n \quad Y_1X_{21} + Y_2X_{22} + \cdots + Y_nX_{2n} \quad \dots \quad Y_1X_{k1} + \\ +Y_2X_{k2} + \cdots + Y_nX_{kn}]$$

Mudando a nomenclatura:

$$\boldsymbol{Y'X} = [a_1 \quad a_2 \quad \dots \quad a_k]$$

Então,

$$\boldsymbol{Y'Xb} = [a_1 \quad a_2 \quad \dots \quad a_k]\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix}$$

$$\boldsymbol{Y'Xb} = a_1b_1 + a_2b_2 + \cdots + a_kb_k$$

Logo,

$$\frac{\partial \boldsymbol{Y'Xb}}{\partial \boldsymbol{b}} = \begin{bmatrix} \frac{\partial \boldsymbol{Y'Xb}}{\partial b_1} \\ \frac{\partial \boldsymbol{Y'Xb}}{\partial b_2} \\ \vdots \\ \frac{\partial \boldsymbol{Y'Xb}}{\partial b_k} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} = (\boldsymbol{Y'X})' = \boldsymbol{X'Y}$$

Portanto,

$$\frac{\partial 2\boldsymbol{Y'Xb}}{\partial \boldsymbol{b}} = 2\boldsymbol{X'Y}$$

A derivada do último termo será dada por:

$$\frac{\partial \boldsymbol{b'X'Xb}}{\partial \boldsymbol{b}} = \cdots$$

Onde

$$\boldsymbol{X'X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{21} & X_{22} & \cdots & X_{2n} \\ X_{31} & X_{32} & \dots & X_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{k1} & X_{k2} & \cdots & X_{kn} \end{bmatrix}\begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \dots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{kn} \end{bmatrix}$$

Ou mudando a nomenclatura

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix} = \boldsymbol{A}$$

Onde $\boldsymbol{A}$ é uma matriz simétrica.

$$\boldsymbol{b}'\boldsymbol{A} = [b_1 \quad b_2 \quad \dots \quad b_k] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$$

Multiplicando

$$\boldsymbol{b}'\boldsymbol{A} = [b_1a_{11} + b_2a_{21} + \cdots + b_ka_{k1} \quad b_1a_{12} + b_2a_{22} + \cdots + b_ka_{k2} \quad \cdots$$
$$b_1a_{1k} + b_2a_{2k} + \cdots + b_ka_{kk}]$$

Que pode-se chamar de

$$\boldsymbol{b}'\boldsymbol{A} = [c_1 \quad c_2 \quad \cdots \quad c_k]$$

Multiplicando pelo vetor restante

$$\boldsymbol{b}'\boldsymbol{A}\boldsymbol{b} = [c_1 \quad c_2 \quad \cdots \quad c_k] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = c_1b_1 + c_2b_2 + \cdots + c_kb_k$$

A derivada será dada por

$$\frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial \boldsymbol{b}} = \begin{bmatrix} \frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial b_1} \\ \frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial b_2} \\ \vdots \\ \frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial b_k} \end{bmatrix}$$

Onde

$$\frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial b_1} = \frac{\partial c_1}{\partial b_1}b_1 + c_1\frac{\partial b_1}{\partial b_1} + \frac{\partial c_2}{\partial b_1}b_2 + c_2\frac{\partial b_2}{\partial b_1} + \cdots + \frac{\partial c_k}{\partial b_1}b_k + c_k\frac{\partial b_k}{\partial b_1}$$

Desenvolvendo

$$\frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial b_1} = a_{11}b_1 + (b_1a_{11} + b_2a_{21} + \cdots + b_ka_{k1}) + a_{12}b_2 + \cdots + a_{1k}b_k$$

$$\frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial b_1} = 2b_1a_{11} + 2b_2a_{12} + \cdots + 2b_ka_{1k}$$

E

$$\frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial b_2} = \frac{\partial c_1}{\partial b_2}b_1 + c_1\frac{\partial b_1}{\partial b_2} + \frac{\partial c_2}{\partial b_2}b_2 + c_2\frac{\partial b_2}{\partial b_2} + \cdots + \frac{\partial c_k}{\partial b_2}b_k + c_k\frac{\partial b_k}{\partial b_2}$$

Ou,

$$\frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial b_2} = a_{21}b_1 + a_{22}b_2 + (b_1a_{12} + b_2a_{22} + \cdots + b_ka_{k2}) + \cdots + a_{2k}b_k$$

$$\frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial b_2} = 2a_{21}b_1 + 2a_{22}b_2 + \cdots + 2a_{2k}b_k$$

Como:

$$\boldsymbol{Ab} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} = \begin{bmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1k}b_k \\ a_{21}b_1 + a_{22}b_2 + \cdots + a_{2k}b_k \\ \vdots \\ a_{k1}b_1 + a_{k2}b_2 + \cdots + a_{kk}b_k \end{bmatrix}$$

Então,

$$\frac{\partial \boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}}{\partial \boldsymbol{b}} = 2\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}$$

Unindo os resultados das derivadas dos termos, tem-se:

$$\frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial \boldsymbol{b}} = \boldsymbol{0} - 2\boldsymbol{X}'\boldsymbol{Y} + 2\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}$$

Como a condição de 1ª ordem é que a equação acima se iguale a zero,

$$\boldsymbol{2X'Xb - 2X'Y = 0}$$

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-\boldsymbol{1}}\boldsymbol{X}'\boldsymbol{Y}$$

13.2. Condição de 2ª Ordem

Além de satisfazer à condição de 1ª ordem deve-se comprovar a condição de 2ª ordem que garante a otimalidade, ou seja, a minimização do erro quadrático. Nesse caso deve-se verificar se a derivada da matriz encontrada para os coeficientes da regressão é positiva definida, conforme enunciado abaixo:

$$\frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial \boldsymbol{b}\partial \boldsymbol{b}'} => Tem\ que\ ser\ positiva\ definida$$

A condição de 1ª ordem permite escrever:

$$\frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial \boldsymbol{b}} = \begin{bmatrix} \frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial b_1} \\ \frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial b_2} \\ \vdots \\ \frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial b_k} \end{bmatrix}$$

Por analogia pode-se afirmar que:

$$\frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial \boldsymbol{b}'} = \begin{bmatrix} \frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial b_1} & \frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial b_2} & \cdots & \frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial b_k} \end{bmatrix}$$

Logo

$$\frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial \boldsymbol{b}'} = \left(\frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial \boldsymbol{b}}\right)' = (-2\boldsymbol{X}'\boldsymbol{Y} + 2\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b})' = -2\boldsymbol{Y}'\boldsymbol{X} + 2\boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}$$

Derivando pela segunda vez em função de $\boldsymbol{b}$, e usando o artifício de derivar um termo de cada vez tem-se:

$$\frac{\partial}{\partial \boldsymbol{b}}(-2\boldsymbol{Y}'\boldsymbol{X}) = \boldsymbol{0}$$

E,

$$\frac{\partial}{\partial \boldsymbol{b}}(2\boldsymbol{b}'\boldsymbol{X}'\boldsymbol{X}) = 2\frac{\partial \boldsymbol{b}'\boldsymbol{A}}{\partial \boldsymbol{b}} = 2\begin{bmatrix} \frac{\partial \boldsymbol{b}'\boldsymbol{A}}{\partial b_1} \\ \frac{\partial \boldsymbol{b}'\boldsymbol{A}}{\partial b_2} \\ \vdots \\ \frac{\partial \boldsymbol{b}'\boldsymbol{A}}{\partial b_k} \end{bmatrix}$$

$$\frac{\partial \boldsymbol{b}'\boldsymbol{A}}{\partial \boldsymbol{b}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix} = \boldsymbol{A}$$

Portanto, somando as derivadas dos termos da equação tem-se:

$$\frac{\partial \boldsymbol{e}'\boldsymbol{e}}{\partial \boldsymbol{b} \partial \boldsymbol{b}'} = \boldsymbol{0} + 2\boldsymbol{X}'\boldsymbol{X}$$

Para provar-se que $\boldsymbol{X}'\boldsymbol{X}$ é positiva definida utiliza-se uma técnica na qual se multiplica a matriz por um vetor e depois pelo mesmo vetor transposto, da seguinte forma:

$$\boldsymbol{v}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{v}$$

Como $\boldsymbol{v}$ possuí dimensões $k \times 1$, $\boldsymbol{X}'\boldsymbol{X}$ apresenta dimensão $k \times k$ e $\boldsymbol{v}'$ $1 \times k$, assim, é possível escrever a equação acima como dois vetores:

$$\boldsymbol{v}'\boldsymbol{X}'\boldsymbol{X}\boldsymbol{v} = \boldsymbol{w}'\boldsymbol{w}$$

O resultado é um escalar igual a:

$$\boldsymbol{w}'\boldsymbol{w} = \sum_{i=1}^{k} \boldsymbol{w_i}^2$$

Desta forma, tem-se que a derivada a segunda do somatório dos resíduos é uma matriz positiva definida. Conclui-se que a condição de 2ª ordem também foi satisfeita e, portanto, o vetor com as melhores estimativas para os coeficientes da regressão será dado pela equação encontrada na condição de primeira ordem.

O desenvolvimento acima comprova que $\boldsymbol{b}$ é um estimador para $\boldsymbol{\beta}$, no entanto, além disso, deseja-se que ele seja o melhor estimador linear não tendencioso (MELNT), ou em inglês *"Best Linear Unbeased Estimator"* (BLUE). A demonstração encontra-se no a seguir.

14
Apêndice E – Demonstração de que $\boldsymbol{b}$ é o MELNT de $\boldsymbol{\beta}$

O primeiro passo é provar que $\boldsymbol{b}$ é um estimador linear. Partindo da equação:

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$$

Demonstrada anteriormente, e da equação da população:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Ambas as equações acima são enunciadas suprimindo o índice das variáveis, o qual indica o número de datas nas quais se dispõem de dados.

Assim como nas suposições do modelo para o erro e para as variáveis, é possível escrever:

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$$

Aplicando a propriedade distributiva:

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}$$

Como,

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X} = \boldsymbol{I}$$

Tem-se:

$$\boldsymbol{b} = \boldsymbol{I}\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}$$

Pode-se representar $\boldsymbol{I}$ e $\boldsymbol{\beta}$ da seguinte maneira:

$$\boldsymbol{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$$

Onde, $\boldsymbol{I}$ possui dimensões $k \times k$ e $\boldsymbol{\beta}$ $k \times 1$.

Portanto,

$$\boldsymbol{I}\boldsymbol{\beta} = \boldsymbol{\beta}$$

Considerando,

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = \boldsymbol{w}$$

Portanto,

$$\boldsymbol{b} = \boldsymbol{\beta} + \boldsymbol{w\varepsilon}$$

Onde, $\boldsymbol{b}$ tem dimensão $k \times 1$, $\boldsymbol{\beta}$ tem dimensão $k \times 1$, $\boldsymbol{w}$ tem dimensão $k \times n$, $\boldsymbol{\varepsilon}$ tem dimensão $n \times 1$ e, portanto, $\boldsymbol{w\varepsilon}$ apresenta dimensão $k \times 1$. Por conta disso, observa-se que $\boldsymbol{b}$ é um estimador linear de $\boldsymbol{\beta}$.

O segundo passo é provar que $\boldsymbol{b}$ é um estimador não tendencioso. A definição de um indicador não tendencioso é que ele não possua um viés de alta ou de baixa. Para que isso ocorra a seguinte relação deve ser atendida:

$$\boldsymbol{E}[\boldsymbol{b}] = \boldsymbol{\beta}$$

Sabe-se que:

$$\boldsymbol{b} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'y}$$

Ou, substituindo $\boldsymbol{y}$ na equação acima pela equação da população,

$$\boldsymbol{b} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'}(\boldsymbol{X\beta} + \boldsymbol{\varepsilon})$$

Aplicando a propriedade distributiva:

$$\boldsymbol{b} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'X\beta} + (\boldsymbol{X'X})^{-1}\boldsymbol{X'\varepsilon}$$

Como,

$$(\boldsymbol{X'X})^{-1}\boldsymbol{X'X} = \boldsymbol{I}$$

Tem-se:

$$\boldsymbol{b} = \boldsymbol{I\beta} + (\boldsymbol{X'X})^{-1}\boldsymbol{X'\varepsilon}$$

Ou,

$$\boldsymbol{b} = \boldsymbol{\beta} + (\boldsymbol{X'X})^{-1}\boldsymbol{X'\varepsilon}$$

Aplicando o valor esperado nos dois lados da equação acima:

$$\boldsymbol{E}[\boldsymbol{b}] = \boldsymbol{E}[\boldsymbol{\beta} + (\boldsymbol{X'X})^{-\boldsymbol{1}}\boldsymbol{X'\varepsilon}]$$

Ou,

$$\boldsymbol{E}[\boldsymbol{b}] = \boldsymbol{E}[\boldsymbol{\beta}] + \boldsymbol{E}[(\boldsymbol{X'X})^{-\boldsymbol{1}}\boldsymbol{X'\varepsilon}]$$

Como a única variável aleatória da equação acima é $\boldsymbol{\varepsilon}$, pode-se escrever:

$$\boldsymbol{E}[\boldsymbol{b}] = \boldsymbol{\beta} + (\boldsymbol{X'X})^{-\boldsymbol{1}}\boldsymbol{X'E}[\boldsymbol{\varepsilon}]$$

No entanto, a variável aleatória acima, que representa o erro de previsão do modelo populacional, foi definida como sendo:

$$\boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$$

O que significa que o valor esperado desta variável será dado por:

$$\boldsymbol{E}[\boldsymbol{\varepsilon}] = \boldsymbol{0}$$

Finalmente, substituindo esse valor esperado do erro aleatório na equação do valor esperado do vetor de coeficientes do modelo, chega-se à equação que se queria demonstrar:

$$\boldsymbol{E}[\boldsymbol{b}] = \boldsymbol{\beta}$$

O terceiro e último passo da demonstração de que $\boldsymbol{b}$ é MELNT de $\boldsymbol{\beta}$, é a de que ele possui a menor variância possível. Neste sentido, inicia-se pela definição de variância:

$$\boldsymbol{Var}[\boldsymbol{b}] = \boldsymbol{E}[(\boldsymbol{b} - \boldsymbol{E}[\boldsymbol{b}])(\boldsymbol{b} - \boldsymbol{E}[\boldsymbol{b}])']$$

Conforme passo anterior é possível escrever:

$$\boldsymbol{b} = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}$$

E,

$$\boldsymbol{E}[\boldsymbol{b}] = \boldsymbol{\beta}$$

Então,

$$\boldsymbol{b} - \boldsymbol{E}[\boldsymbol{b}] = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon} - \boldsymbol{\beta}$$

Ou,

$$\boldsymbol{b} - \boldsymbol{E}[\boldsymbol{b}] = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon}$$

Logo,

$$\boldsymbol{Var}[\boldsymbol{b}] = \boldsymbol{E}[((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon})((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon})']$$

Ou,

$$\boldsymbol{Var}[\boldsymbol{b}] = \boldsymbol{E}[((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\varepsilon})\boldsymbol{\varepsilon}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}]$$

Como

$$[(\boldsymbol{X}'\boldsymbol{X})^{-1}]' = (\boldsymbol{X}'\boldsymbol{X})^{-1}$$

Como só há duas variáveis aleatórias na equação acima, pode-se escrever:

$$\boldsymbol{Var}[\boldsymbol{b}] = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}']\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

Mas sabe-se que outra forma de se escrever os vetores de erro aleatório populacional é:

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

E, portanto,

$$\boldsymbol{\varepsilon}' = [\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n]$$

Onde, $\boldsymbol{\varepsilon}$ possui dimensões $n \times 1$, e consequentemente $\boldsymbol{\varepsilon}'$ possui dimensões $1 \times n$. Neste sentido, o produto dessas variáveis será dado por:

$$\boldsymbol{\varepsilon\varepsilon}' = \begin{bmatrix} {\varepsilon_1}^2 & \varepsilon_1\varepsilon_2 & \cdots & \varepsilon_1\varepsilon_n \\ \varepsilon_2\varepsilon_1 & {\varepsilon_2}^2 & \cdots & \varepsilon_2\varepsilon_n \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_n\varepsilon_1 & \varepsilon_n\varepsilon_2 & \cdots & {\varepsilon_n}^2 \end{bmatrix}$$

Ao aplicar-se o valor esperado em todos os termos da matriz será possível constatar que:

$$E[{\varepsilon_i}^2] = Var[\varepsilon_i]$$

Onde a variância de ε_i é dada pela definição, e é igual a σ^2.

$$E\left[\varepsilon_i\varepsilon_j\right] = cov\left[\varepsilon_i\varepsilon_j\right]$$

Mas a covariância entre os erros aleatórios é definida como zero pelo modelo. Desta forma, a matriz de variâncias-covariâncias será dada por:

$$\boldsymbol{E}[\boldsymbol{\varepsilon\varepsilon}'] = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}$$

Ou,

$$\boldsymbol{E}[\boldsymbol{\varepsilon\varepsilon}'] = \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Escrevendo de outra maneira:

$$\boldsymbol{E}[\boldsymbol{\varepsilon\varepsilon}'] = \sigma^2\boldsymbol{I}$$

Substituindo o valor encontrado na equação de variância dos coeficientes do modelo amostral:

$$\boldsymbol{Var}[\boldsymbol{b}] = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\sigma^2\boldsymbol{I}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

Desenvolvendo,

$$\boldsymbol{Var}[\boldsymbol{b}] = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

Que é igual a:

$$\boldsymbol{Var}[\boldsymbol{b}] = \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

Ou, alterando as variáveis de tal forma que:

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = \boldsymbol{W}$$

E, portanto,

$$\boldsymbol{W}' = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

Assim,

$$\boldsymbol{Var}[\boldsymbol{b}] = \sigma^2\boldsymbol{W}\boldsymbol{W}'$$

Como,

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$$

Então:

$$\boldsymbol{b} = \boldsymbol{W}\boldsymbol{y}$$

Considerando a estimação de um novo vetor de coeficientes, da seguinte maneira:

$$\widehat{\boldsymbol{b}} = \boldsymbol{G}\boldsymbol{y}$$

Substituindo a equação da população na relação acima se chega a:

$$\widehat{\boldsymbol{b}} = \boldsymbol{G}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{G}\boldsymbol{\varepsilon}$$

Da mesma forma que se demonstrou para o vetor $\boldsymbol{b}$ é possível se provar que para que:

$$\boldsymbol{E}[\widehat{\boldsymbol{b}}] = \boldsymbol{\beta}$$

Algumas propriedades devem ser atendidas, são elas:

$$\boldsymbol{E}[\widehat{\boldsymbol{b}}] = \boldsymbol{E}[\boldsymbol{G}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{G}\boldsymbol{\varepsilon}]$$

Aplicando a propriedade da soma dos valores esperados:

$$\boldsymbol{E}[\widehat{\boldsymbol{b}}] = \boldsymbol{E}[\boldsymbol{G}\boldsymbol{X}\boldsymbol{\beta}] + \boldsymbol{E}[\boldsymbol{G}\boldsymbol{\varepsilon}]$$

Como só há uma variável aleatória na equação acima, pode-se operar com o valor esperado da seguinte forma:

$$\boldsymbol{E}[\widehat{\boldsymbol{b}}] = \boldsymbol{G}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{G}\boldsymbol{E}[\boldsymbol{\varepsilon}]$$

Como o valor esperado do erro aleatório populacional é igual a zero chega-se ao resultado:

$$\boldsymbol{E}[\widehat{\boldsymbol{b}}] = \boldsymbol{G}\boldsymbol{X}\boldsymbol{\beta}$$

Mas, o objetivo é que $\widehat{\boldsymbol{b}}$ seja um estimador não viesado de $\boldsymbol{\beta}$, e para tanto a seguinte igualdade deve ser satisfeita:

$$\boldsymbol{G}\boldsymbol{X} = \boldsymbol{I}$$

O próximo passo é obter a variância do estimador $\widehat{\boldsymbol{b}}$,

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \boldsymbol{E}\left[\left(\widehat{\boldsymbol{b}} - \boldsymbol{E}[\widehat{\boldsymbol{b}}]\right)\left(\widehat{\boldsymbol{b}} - \boldsymbol{E}[\widehat{\boldsymbol{b}}]\right)'\right]$$

Onde,

$$\widehat{\boldsymbol{b}} - \boldsymbol{E}[\widehat{\boldsymbol{b}}] = \widehat{\boldsymbol{b}} = \boldsymbol{G}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{G}\boldsymbol{\varepsilon} - \boldsymbol{G}\boldsymbol{X}\boldsymbol{\beta}$$

Simplificando,

$$\widehat{\boldsymbol{b}} - \boldsymbol{E}[\widehat{\boldsymbol{b}}] = \widehat{\boldsymbol{b}} = \boldsymbol{I}\boldsymbol{\beta} + \boldsymbol{G}\boldsymbol{\varepsilon} - \boldsymbol{I}\boldsymbol{\beta}$$

Ou,

$$\widehat{\boldsymbol{b}} - \boldsymbol{E}[\widehat{\boldsymbol{b}}] = \widehat{\boldsymbol{b}} = \boldsymbol{G}\boldsymbol{\varepsilon}$$

Portanto,

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \boldsymbol{E}[(\boldsymbol{G\varepsilon})(\boldsymbol{G\varepsilon})']$$

Aplicando a propriedade de transposição da multiplicação de duas matrizes:

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \boldsymbol{E}[(\boldsymbol{G\varepsilon})\boldsymbol{\varepsilon}'\boldsymbol{G}']$$

Entrando com o valor esperado,

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \boldsymbol{GE}[\boldsymbol{\varepsilon\varepsilon}']\boldsymbol{G}'$$

Substituindo a matriz de variâncias-covariâncias dos erros populacionais,

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \boldsymbol{G}\sigma^2\boldsymbol{IG}'$$

Ou,

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \sigma^2\boldsymbol{GG}'$$

Agora o objetivo passa a ser provar que a variância do estimador $\boldsymbol{b}$ é menor do que a do estimador $\widehat{\boldsymbol{b}}$. Essa prova pode ser feita demonstrando que:

$$\boldsymbol{G} = \boldsymbol{W} + \boldsymbol{D}$$

Pois, as matrizes de variâncias de $\boldsymbol{b}$ e de $\widehat{\boldsymbol{b}}$ são dadas por:

$$\boldsymbol{Var}[\boldsymbol{b}] = \sigma^2\boldsymbol{WW}'$$

E,

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \sigma^2\boldsymbol{GG}'$$

Respectivamente. A seguir escreve-se a matriz de variâncias de $\widehat{\boldsymbol{b}}$ em função de $\boldsymbol{W}$ e $\boldsymbol{D}$.

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \sigma^2(\boldsymbol{W} + \boldsymbol{D})\ (\boldsymbol{W} + \boldsymbol{D})'$$

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \sigma^2(\boldsymbol{W} + \boldsymbol{D})\ \boldsymbol{W}' + \boldsymbol{D}'$$

Aplicando a propriedade distributiva,

$$\boldsymbol{Var}[\widehat{\boldsymbol{b}}] = \sigma^2(\boldsymbol{WW}' + \boldsymbol{WD}' + \boldsymbol{DW}' + \boldsymbol{DD}')$$

Como,

$$\boldsymbol{GX} = \boldsymbol{I}$$

Então,

$$(\boldsymbol{W} + \boldsymbol{D})\boldsymbol{X} = \boldsymbol{I}$$

Abrindo,

$$\boldsymbol{WX} + \boldsymbol{DX} = \boldsymbol{I}$$

Definiu-se que,

$$\boldsymbol{W} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$$

Substituindo $\boldsymbol{W}$ na equação anterior,

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{DX} = \boldsymbol{I}$$

Portanto,

$$\boldsymbol{I} + \boldsymbol{DX} = \boldsymbol{I}$$

Logo,

$$\boldsymbol{DX} = \boldsymbol{I} - \boldsymbol{I}$$

Consequentemente,

$$\boldsymbol{DX} = \boldsymbol{0}$$

Da mesma forma,

$$(\boldsymbol{DX})' = \boldsymbol{0}'$$

Então,

$$\boldsymbol{X}'\boldsymbol{D}' = \boldsymbol{0}$$

Substituindo $\boldsymbol{W}$ em $\boldsymbol{WD}'$,

$$\boldsymbol{WD}' = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{D}'$$

E substituindo $\boldsymbol{W}'$ em $\boldsymbol{DW}'$,

$$\boldsymbol{DW}' = \boldsymbol{DX}(\boldsymbol{X}'\boldsymbol{X})^{-1}$$

Mas, como encontrado anteriormente, tanto $\boldsymbol{X}'\boldsymbol{D}'$ quanto $\boldsymbol{DX}$, são iguais a zero, e, portanto:

$$\boldsymbol{WD}' = \boldsymbol{0}$$

E,

$$\boldsymbol{DW}' = \boldsymbol{0}$$

Assim,

$$\boldsymbol{Var}[\hat{\boldsymbol{b}}] = \sigma^2(\boldsymbol{WW}' + \boldsymbol{DD}')$$

Explicitando a matriz $\boldsymbol{W}$,

$$\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{k1} & w_{k2} & \cdots & w_{kn} \end{bmatrix}$$

Logo,

$$\boldsymbol{W}' = \begin{bmatrix} w_{11} & w_{21} & \cdots & w_{k1} \\ w_{12} & w_{22} & \cdots & w_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1n} & w_{2n} & \cdots & w_{kn} \end{bmatrix}$$

O produto das duas matrizes será dado por:

$$\boldsymbol{WW'} = \begin{bmatrix} \sum_{i=1}^{n} w_{1i}^{2} & \sum_{i=1}^{n} w_{1i}w_{2i} & \cdots & \sum_{i=1}^{n} w_{1i}w_{ki} \\ \sum_{i=1}^{n} w_{2i}w_{1i} & \sum_{i=1}^{n} w_{2i}^{2} & \cdots & \sum_{i=1}^{n} w_{2i}w_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} w_{ki}w_{1i} & \sum_{i=1}^{n} w_{ki}w_{2i} & \cdots & \sum_{i=1}^{n} w_{ki}^{2} \end{bmatrix}$$

Se for incluído σ^2 na equação acima, a matriz passa a ser a de variâncias-covariâncias e pode ser escrita como:

$$\sigma^2\boldsymbol{ww'} = \begin{bmatrix} Var[b_1] & Cov[b_1, b_2] & \cdots & Cov[b_1, b_k] \\ Cov[b_2, b_1] & Var[b_2] & \cdots & Cov[b_2, b_k] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[b_k, b_1] & Cov[b_k, b_2] & \cdots & Var[b_k] \end{bmatrix}$$

As variâncias de $\boldsymbol{\hat{b}}$ serão dadas por:

$$Var\left[\widehat{b_1}\right] = \left(\sum_{i=1}^{n} w_{1i}^{2} + \sum_{i=1}^{n} d_{1i}^{2}\right)\sigma^2$$

$$Var\left[\widehat{b_2}\right] = \left(\sum_{i=1}^{n} w_{2i}^{2} + \sum_{i=1}^{n} d_{2i}^{2}\right)\sigma^2$$

Até,

$$Var\left[\widehat{b_k}\right] = \left(\sum_{i=1}^{n} w_{ki}^{2} + \sum_{i=1}^{n} d_{ki}^{2}\right)\sigma^2$$

Comparando as variâncias acima com as do vetor $\boldsymbol{b}$ é possível constatar que:

$$Var\left[\widehat{b_1}\right] = Var[b_1] + \sigma^2 \sum_{i=1}^{n} d_{1i}^{2}$$

$$Var\left[\widehat{b_2}\right] = Var[b_2] + \sigma^2 \sum_{i=1}^{n} d_{2i}^{2}$$

Até,

$$Var\left[\widehat{b_k}\right] = Var[b_k] + \sigma^2 \sum_{i=1}^{n} d_{ki}^{2}$$

Como a todas as variâncias dos coeficientes estimados em $\boldsymbol{\hat{b}}$ é somado um termo elevado ao quadrado, que sempre será maior ou igual a zero, é possível concluir que as variâncias dos coeficientes de $\boldsymbol{\hat{b}}$ serão sempre maiores ou iguais as de $\boldsymbol{b}$. Como a

$$\boldsymbol{Var}\left[\boldsymbol{\hat{b}}\right] \geq \boldsymbol{Var}[\boldsymbol{b}]$$

Conclui-se que $\boldsymbol{b}$ é MELNT de $\boldsymbol{\beta}$.

15
Apêndice F – Demonstração do Algoritmo Recursivo de MQO

Abaixo é apresentada uma maneira de se encontrar o vetor de coeficientes da regressão em $t+1$ em função do vetor de t.

15.1. Demonstração de $\boldsymbol{b_{t+1}}$

Deseja-se provar a equação de atualização do vetor de coeficientes da regressão, tal que

$$\boldsymbol{b_{t+1}} = \boldsymbol{b_t} + \boldsymbol{K_{t+1}}(\boldsymbol{Y_{t+1}} - \boldsymbol{x'_{t+1}}\boldsymbol{b_t})$$

A equação acima pode ser obtida através das seguintes relações:

$$\boldsymbol{b_t} = (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{Y_t}$$

Então,

$$\boldsymbol{b_{t+1}} = (\boldsymbol{X'_{t+1}}\boldsymbol{X_{t+1}})^{-1}\boldsymbol{X'_{t+1}}\boldsymbol{Y_{t+1}}$$

$$\boldsymbol{b_{t+1}} = (\boldsymbol{X'_{t+1}}\boldsymbol{X_{t+1}})^{-\mathbf{1}}[\boldsymbol{X'_t} \quad \boldsymbol{x_{t+1}}]\begin{bmatrix}\boldsymbol{Y_t}\\ \boldsymbol{Y_{t+1}}\end{bmatrix}$$

Onde:

$$\boldsymbol{x_{t+1}} = \begin{bmatrix} 1 \\ X_{2t+1} \\ X_{3t+1} \\ \vdots \\ X_{kt+1} \end{bmatrix}$$

$$\boldsymbol{X_{t+1}} = \begin{bmatrix}\boldsymbol{X_t}\\ \boldsymbol{x'_{t+1}}\end{bmatrix}$$

$$\boldsymbol{X_{t+1}} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{2t+1} & X_{3t+1} & \cdots & X_{kt+1} \end{bmatrix}$$

$$\boldsymbol{X'_{t+1}} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{21} & X_{22} & \cdots & X_{2t+1} \\ X_{31} & X_{32} & & X_{3t+1} \\ \vdots & \vdots & \ddots & \vdots \\ X_{k1} & X_{k2} & \cdots & X_{kt+1} \end{bmatrix}$$

E o índice t corresponde ao n das condições de primeira e segunda ordens.

Consequentemente,

$$\boldsymbol{b}_{t+1} = (\boldsymbol{X}'_{t+1}\boldsymbol{X}_{t+1})^{-1}[\boldsymbol{X}'_t\boldsymbol{Y}_t + \boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1}]$$

$$\boldsymbol{b}_{t+1} = (\boldsymbol{X}'_{t+1}\boldsymbol{X}_{t+1})^{-1}\boldsymbol{X}'_t\boldsymbol{Y}_t + (\boldsymbol{X}'_{t+1}\boldsymbol{X}_{t+1})^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1}$$

Ou,

$$\boldsymbol{b}_{t+1} = \boldsymbol{F} + \boldsymbol{G}$$

Pode-se escrever, ainda:

$$\boldsymbol{X}'_{t+1}\boldsymbol{X}_{t+1} = [\boldsymbol{X}'_t \quad \boldsymbol{x}_{t+1}]\begin{bmatrix}\boldsymbol{X}_t \\ \boldsymbol{x}'_{t+1}\end{bmatrix}$$

$$\boldsymbol{X}'_{t+1}\boldsymbol{X}_{t+1} = \boldsymbol{X}'_t\boldsymbol{X}_t + \boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}$$

Para prosseguir com o raciocínio é preciso utilizar o Lema de Inversão:

$$(\boldsymbol{A} + \boldsymbol{BC})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I} + \boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CA}^{-1}$$

Esse Lema pode ser facilmente demonstrado, e a demonstração está no Apêndice H.

Deseja-se obter a inversa da equação encontrada acima, ou seja,

$$(\boldsymbol{X}'_{t+1}\boldsymbol{X}_{t+1})^{-1} = (\boldsymbol{X}'_t\boldsymbol{X}_t + \boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1})^{-1}$$

Aplicando o Lema de Inversão de matrizes se obtém:

$$(\boldsymbol{X}'_t\boldsymbol{X}_t + \boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1})^{-1} = (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1} + (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}(\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1})^{-1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}$$

Onde:

$$\boldsymbol{X}'_t\boldsymbol{X}_t = \boldsymbol{A}$$

$$\boldsymbol{x}_{t+1} = \boldsymbol{B}$$

$$\boldsymbol{x}'_{t+1} = \boldsymbol{C}$$

Como,

$$\boldsymbol{F} = (\boldsymbol{X}'_{t+1}\boldsymbol{X}_{t+1})^{-1}\boldsymbol{X}'_t\boldsymbol{Y}_t$$

Então,

$$\boldsymbol{F} = \left((\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1} - (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\left(\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\right)^{-1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\right)\boldsymbol{X}'_t\boldsymbol{Y}_t$$

A dimensão do termo $(\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1})^{-1}$ é 1×1, pois como $\boldsymbol{X}_t$ é $t \times k$ e, consequentemente, $\boldsymbol{X}'_t$ é $k \times t$, e o produto entre eles é $k \times k$. Da mesma forma, como $\boldsymbol{x}'_{t+1}$ representa apenas uma nova observação das variáveis independentes, e, portanto, possui dimensão $1 \times k$, o produto $\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}$ também apresenta dimensão $1 \times k$. $\boldsymbol{x}_{t+1}$ possui dimensão $k \times 1$, logo

$\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}$ terá dimensão 1×1, conforme afirmado acima. Considerando esse termo como um escalar, é possível escrever:

$$\boldsymbol{F} = \left((\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1} - \frac{(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}}{\boldsymbol{I}+\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}\right)\boldsymbol{X}'_t\boldsymbol{Y}_t$$

$$\boldsymbol{F} = (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{X}'_t\boldsymbol{Y}_t - \frac{(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{X}'_t\boldsymbol{Y}_t}{\boldsymbol{I}+\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}$$

$$\boldsymbol{F} = \boldsymbol{b}_t - \frac{(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}\boldsymbol{b}_t}{\boldsymbol{I}+\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}$$

Considerando:

$$\boldsymbol{K}_{t+1} = \frac{(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}{\boldsymbol{I}+\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}$$

Obtêm-se:

$$\boldsymbol{F} = \boldsymbol{b}_t - \boldsymbol{K}_{t+1}\boldsymbol{x}'_{t+1}\boldsymbol{b}_t$$

O outro termo da equação de atualização foi definido como:

$$\boldsymbol{G} = (\boldsymbol{X}'_{t+1}\boldsymbol{X}_{t+1})^{-1}\boldsymbol{X}_{t+1}\boldsymbol{Y}_{t+1}$$

Aplicando o resultado encontrado com o Lema de Inversão,

$$\boldsymbol{G} = ((\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1} - (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}(\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1})^{-1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1})\ \boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1}$$

$$\boldsymbol{G} = \left((\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1} - \frac{(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}}{\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}\right)\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1}$$

$$\boldsymbol{G} = (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1} - \frac{(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1}}{\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}$$

$$\boldsymbol{G} = \frac{1}{\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}((\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1}(\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}) - (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1})$$

$$\boldsymbol{G} = \frac{\boldsymbol{1}}{\boldsymbol{D}}((\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1}(\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}) + (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1})$$

Onde:

$$\boldsymbol{D} = \boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}$$

Então,

$$\boldsymbol{G} = \frac{1}{\boldsymbol{D}}((\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1} + (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1} + -(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1})$$

Como $\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}$ possui dimensão 1×1 pode-se escrever:

$$\boldsymbol{G} = \frac{\boldsymbol{1}}{\boldsymbol{D}}((\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1} + (\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1} + -(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1})$$

Portanto, como os dois últimos termos são exatamente iguais,

$$\boldsymbol{G} = \frac{\boldsymbol{1}}{\boldsymbol{D}}((\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1})$$

$$\boldsymbol{G} = \frac{(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}\boldsymbol{Y}_{t+1}}{\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}$$

Como,

$$\boldsymbol{K}_{t+1} = \frac{(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}{\boldsymbol{I} + \boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}$$

Obtêm-se:

$$\boldsymbol{G} = \boldsymbol{K}_{t+1}\boldsymbol{Y}_{t+1}$$

Definiu-se:

$$\boldsymbol{b}_{t+1} = \boldsymbol{F} + \boldsymbol{G}$$

Então, substituindo $\boldsymbol{F}$ e $\boldsymbol{G}$ da equação acima pelas relações encontradas nas equações anteriores obtêm-se:

$$\boldsymbol{b}_{t+1} = \boldsymbol{b}_t - \boldsymbol{K}_{t+1}\boldsymbol{x}'_{t+1}\boldsymbol{b}_t + \boldsymbol{K}_{t+1}\boldsymbol{Y}_{t+1}$$

Ou,

$$\boldsymbol{b}_{t+1} = \boldsymbol{b}_t + \boldsymbol{K}_{t+1}(\boldsymbol{Y}_{t+1} - \boldsymbol{x}'_{t+1}\boldsymbol{b}_t)$$

Desta forma, ficou demonstrada a equação, que utiliza o vetor de coeficientes $\boldsymbol{b}_t$ no tempo t para obter uma nova estimativa para vetor de coeficientes em $t + 1$. A estimativa atualizada é obtida tanto através da antiga quanto do erro de previsão associado a um fator de correção $\boldsymbol{K}_t$.

No caso da estimação por mínimos quadrados ordinários, como foi citado acima,

$$\boldsymbol{K}_{t+1} = \frac{(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}{\boldsymbol{I}+\boldsymbol{x}'_{t+1}(\boldsymbol{X}'_t\boldsymbol{X}_t)^{-1}\boldsymbol{x}_{t+1}}$$

E,

$$\boldsymbol{b}_{t+1} = \boldsymbol{b}_t + \boldsymbol{K}_{t+1}(\boldsymbol{Y}_{t+1} - \boldsymbol{x}'_{t+1}\boldsymbol{b}_t)$$

As duas equações acima formam um algoritmo recursivo para atualizar estimativas.

15.2. Demonstração de $\boldsymbol{P_{t+1}}$

Agora deseja-se simplificar o algoritmo criado acima. Nesse sentido opta-se por alterar a nomenclatura utilizada da seguinte forma:

$$\boldsymbol{P_t} = (\boldsymbol{X_t'X_t})^{-1}$$

Então,

$$\boldsymbol{K_{t+1}} = \frac{\boldsymbol{P_t x_{t+1}}}{\boldsymbol{I + x_{t+1}' P_t x_{t+1}}}$$

A mudança realizada apresenta como vantagem o fato de não ser mais preciso inverter matrizes de grandes dimensões, como pode ser o caso de $(\boldsymbol{X_t'X_t})^{-1}$, que apresenta dimensões $k \times k$, que podem ser tão maiores quanto mais complexo for o modelo adotado, ou seja, quanto maior for o número de variáveis explicativas ou independentes. O artifício da eliminação da necessidade de se inverter a matriz é alcançado através da atualização de $\boldsymbol{P_t}$, que é feita seguindo a relação:

$$\boldsymbol{P_{t+1}} = \boldsymbol{P_t} - \boldsymbol{P_t} \frac{\boldsymbol{x_{t+1}x_{t+1}'}}{\boldsymbol{I + x_{t+1}' P_t x_{t+1}}} \boldsymbol{P_t}$$

A atualização da equação pode ser facilmente demonstrada utilizando o Lema de Inversão de matrizes:

$$(\boldsymbol{A} + \boldsymbol{BC})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I} + \boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CA}^{-1}$$

Como,

$$\boldsymbol{P_t} = (\boldsymbol{X_t'X_t})^{-1}$$

Então,

$$\boldsymbol{P_{t+1}} = (\boldsymbol{X_{t+1}'X_{t+1}})^{-1}$$

$$\boldsymbol{P_{t+1}} = (\boldsymbol{X_t'X_t} + \boldsymbol{x_{t+1}x_{t+1}'})^{-1}$$

Aplicando o Lema de Inversão de matrizes na equação acima:

$$\boldsymbol{P_{t+1}} = (\boldsymbol{X_t'X_t})^{-1} + (\boldsymbol{X_t'X_t})^{-1}\boldsymbol{x_{t+1}}(\boldsymbol{I} + \boldsymbol{x_{t+1}'}(\boldsymbol{X_t'X_t})^{-1}\boldsymbol{x_{t+1}})^{-1}\boldsymbol{x_{t+1}'}(\boldsymbol{X_t'X_t})^{-1}$$

Onde:

$$\boldsymbol{X_t'X_t} = \boldsymbol{A}$$

$$\boldsymbol{x_{t+1}} = \boldsymbol{B}$$

$$\boldsymbol{x_{t+1}'} = \boldsymbol{C}$$

Como o termo $\boldsymbol{I} + \boldsymbol{x_{t+1}'}(\boldsymbol{X_t'X_t})^{-1}\boldsymbol{x_{t+1}}$ apresenta dimensões 1×1 é possível escrever:

$$\boldsymbol{P_{t+1}} = (\boldsymbol{X_t'X_t})^{-1} - (\boldsymbol{X_t'X_t})^{-1}\frac{\boldsymbol{x_{t+1}x_{t+1}'}}{\boldsymbol{I}+\boldsymbol{x_{t+1}'}(\boldsymbol{X_t'X_t})^{-1}\boldsymbol{x_{t+1}}}(\boldsymbol{X_t'X_t})^{-1}$$

Ou, substituindo a equação de $\boldsymbol{P_t}$ na de $\boldsymbol{P_{t+1}}$,

$$\boldsymbol{P_{t+1}} = \boldsymbol{P_t} - \boldsymbol{P_t}\frac{\boldsymbol{x_{t+1}x_{t+1}'}}{\boldsymbol{I} + \boldsymbol{x_{t+1}'P_t x_{t+1}}}\boldsymbol{P_t}$$

Assim, chega-se a equação de atualização de $\boldsymbol{P_t}$.

15.3. Relação Entre a Variância de $\boldsymbol{b_t}$ e $\boldsymbol{P_t}$

A matriz de covariâncias da regressão dos coeficientes de $\boldsymbol{b_t}$ é dada por:

$$\boldsymbol{Var}(\boldsymbol{b_t}) = \sigma^2\boldsymbol{P_t}$$

A equação acima pode ser obtida partindo da equação dos coeficientes,

$$\boldsymbol{b_t} = (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{Y_t}$$

Como,

$$\boldsymbol{Y_t} = \boldsymbol{X_t\beta_t} + \boldsymbol{\varepsilon_t}$$

Então,

$$\boldsymbol{b_t} = (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'(\boldsymbol{X_t\beta_t} + \boldsymbol{\varepsilon_t})$$

$$\boldsymbol{b_t} = (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{X_t\beta_t} + (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{\varepsilon_t}$$

$$\boldsymbol{b_t} = \boldsymbol{\beta_t} + (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{\varepsilon_t}$$

Subtraindo $\boldsymbol{\beta_t}$ dos dois lados da equação:

$$\boldsymbol{b_t} - \boldsymbol{\beta_t} = \boldsymbol{\beta_t} + (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{\varepsilon_t} - \boldsymbol{\beta_t}$$

Do Apêndice E sabe-se que:

$$\boldsymbol{E}[\boldsymbol{b_t}] = \boldsymbol{\beta_t}$$

Substituindo,

$$\boldsymbol{b_t} - \boldsymbol{E}[\boldsymbol{b_t}] = \boldsymbol{\beta_t} + (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{\varepsilon_t} - \boldsymbol{\beta_t}$$

$$\boldsymbol{b_t} - \boldsymbol{E}[\boldsymbol{b_t}] = (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{\varepsilon_t}$$

Uma das definições para variância é:

$$\boldsymbol{Var}(\boldsymbol{b_t}) = \boldsymbol{E}[(\boldsymbol{b_t} - \boldsymbol{E}[\boldsymbol{b_t}])^{\boldsymbol{2}}]$$

Portanto,

$$\boldsymbol{Var}(\boldsymbol{b_t}) = E[((\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{\varepsilon_t})((\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{\varepsilon_t})']$$

$$\boldsymbol{Var}(\boldsymbol{b}) = \boldsymbol{E}[((\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{\varepsilon_t})\boldsymbol{\varepsilon'_t X_t}(\boldsymbol{X'_t X_t})^{\boldsymbol{-1}}]$$

Como a única variável aleatória é $\boldsymbol{\varepsilon_t}$,

$$\boldsymbol{Var}(\boldsymbol{b_t}) = (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{E}[\boldsymbol{\varepsilon_t\varepsilon'_t}]\boldsymbol{X_t}(\boldsymbol{X'_t X_t})^{\boldsymbol{-1}}$$

A variável $\boldsymbol{\varepsilon_t}$ foi definida como:

$\boldsymbol{\varepsilon_t} \sim \boldsymbol{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$

Mas,

$\boldsymbol{Var}(\boldsymbol{\varepsilon_t}) = \boldsymbol{E}[(\boldsymbol{\varepsilon_t} - \boldsymbol{E}[\boldsymbol{\varepsilon_t}])^{\boldsymbol{2}}]$

$\boldsymbol{Var}(\boldsymbol{\varepsilon_t}) = \boldsymbol{E}[(\boldsymbol{\varepsilon_t} - \boldsymbol{E}[\boldsymbol{\varepsilon_t}])(\boldsymbol{\varepsilon_t} - \boldsymbol{E}[\boldsymbol{\varepsilon_t}])']$

Como,

$\boldsymbol{E}[\boldsymbol{\varepsilon_t}] = \boldsymbol{0}$

Então,

$\boldsymbol{Var}(\boldsymbol{\varepsilon_t}) = \boldsymbol{E}[\boldsymbol{\varepsilon_t}\boldsymbol{\varepsilon'_t}]$

Mas pela definição,

$\boldsymbol{Var}(\boldsymbol{\varepsilon_t}) = \sigma^2 \boldsymbol{I}$

Portanto,

$\boldsymbol{Var}(\boldsymbol{b_t}) = (\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\sigma^2 \boldsymbol{I}\boldsymbol{X_t}(\boldsymbol{X_t}'\boldsymbol{X_t})^{\boldsymbol{-1}}$

Ou,

$\boldsymbol{Var}(\boldsymbol{b_t}) = \sigma^2(\boldsymbol{X_t}'\boldsymbol{X_t})^{-1}\boldsymbol{X_t}'\boldsymbol{X_t}(\boldsymbol{X_t}'\boldsymbol{X_t})^{\boldsymbol{-1}}$

Consequentemente,

$\boldsymbol{Var}(\boldsymbol{b_t}) = \sigma^2(\boldsymbol{X_t}'\boldsymbol{X_t})^{\boldsymbol{-1}}$

Ou,

$\boldsymbol{Var}(\boldsymbol{b_t}) = \sigma^2 \boldsymbol{P_t}$

Logo,

$\boldsymbol{Var}(\boldsymbol{b_{t+1}}) = \sigma^2 \boldsymbol{P_{t+1}}$

15.4. Resumo das Equações do Algoritmo Recursivo do MQO

As equações,

$\boldsymbol{b_{t+1}} = \boldsymbol{b_t} + \boldsymbol{K_{t+1}}(\boldsymbol{Y_{t+1}} - \boldsymbol{x'_{t+1}}\boldsymbol{b_t})$

$\boldsymbol{K_{t+1}} = \frac{(\boldsymbol{X'_t X_t})^{-1}\boldsymbol{x_{t+1}}}{\boldsymbol{I} + \boldsymbol{x'_{t+1}}(\boldsymbol{X'_t X_t})^{-1}\boldsymbol{x_{t+1}}}$

$\boldsymbol{P_t} = (\boldsymbol{X'_t X_t})^{\boldsymbol{-1}}$

$\boldsymbol{K_{t+1}} = \frac{\boldsymbol{P_t x_{t+1}}}{\boldsymbol{I} + \boldsymbol{x'_{t+1}P_t x_{t+1}}}$

$\boldsymbol{P_{t+1}} = \boldsymbol{P_t} - \boldsymbol{P_t}\frac{\boldsymbol{x_{t+1}x'_{t+1}}}{\boldsymbol{I} + \boldsymbol{x'_{t+1}P_t x_{t+1}}}\boldsymbol{P_t}$

Formam um algoritmo recursivo para o computo do parâmetro $\boldsymbol{b_t}$.

16
Apêndice G – Demonstrações do Filtro de Kalman

A seguir são apresentadas as demonstrações referentes ao algoritmo do FK.

16.1. Transformação das Equações do FK Para Equações com Constantes

Algumas vezes as equações do Filtro de Kalman aparecem com constantes. Cabe ressaltar que as equações apresentadas neste texto incorporam essas constantes de forma implícita. A equação de medida pode ser escrita como

$$\boldsymbol{z_t} = \boldsymbol{H_t}\boldsymbol{x_t} + \boldsymbol{u_t}$$

Ou abrindo,

$$\boldsymbol{z_t} = [\boldsymbol{d_t} \quad \boldsymbol{F_t}] \begin{bmatrix} 1 \\ \boldsymbol{w_t} \end{bmatrix} + \boldsymbol{u_t}$$

Multiplicando os termos das matrizes acima

$$\boldsymbol{z_t} = \boldsymbol{d_t} + \boldsymbol{F_t}\boldsymbol{w_t} + \boldsymbol{u_t}$$

Onde:

$$\boldsymbol{d_t} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \text{; com dimensão } n \times 1.$$

$$\boldsymbol{F_t} = \begin{bmatrix} F_{12} & F_{13} & \dots & F_{1k} \\ F_{22} & F_{23} & \dots & F_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ F_{n1} & F_{n2} & \dots & F_{nk} \end{bmatrix} \text{; com dimensão } n \times k - 1.$$

$$\boldsymbol{w_t} = \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix} \text{; com dimensão } k - 1 \times 1$$

1 representa um escalar.

O mesmo pode ser feito para a equação de transição que também incorpora constantes implicitamente.

$$\boldsymbol{x_t} = \boldsymbol{\phi}\boldsymbol{x_{t-1}} + \boldsymbol{G}\boldsymbol{v_t}$$

Que pode ser escrito como

$$\boldsymbol{x_t} = [\boldsymbol{c_t} \quad \boldsymbol{B}]\begin{bmatrix} 1 \\ \boldsymbol{w_{t-1}} \end{bmatrix} + \boldsymbol{G}\boldsymbol{v_t}$$

Ou,

$$\boldsymbol{x_t} = \boldsymbol{c_t} + \boldsymbol{B}\boldsymbol{w_{t-1}} + \boldsymbol{G}\boldsymbol{v_t}$$

Onde:

$$\boldsymbol{B} = \begin{bmatrix} B_{12} & B_{13} & \dots & B_{1k} \\ B_{22} & B_{23} & \dots & B_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ B_{k2} & B_{k3} & \dots & B_{kk} \end{bmatrix}; \text{ com dimensão } k \times k-1.$$

$$\boldsymbol{w_{t-1}} = \begin{bmatrix} x_{2t-1} \\ x_{3t-1} \\ \vdots \\ x_{kt-1} \end{bmatrix}; \text{ com dimensão } k-1 \times 1.$$

$$\boldsymbol{c_t} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}; \text{ com dimensão } k \times 1.$$

$$\boldsymbol{G} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}; \text{ com dimensão } k \times k.$$

1 representa um escalar.

Assim, ficou demonstrado que as equações do FK podem ser escritas de duas formas, explicitando as constantes ou não, e que elas são exatamente iguais.

16.2. Demonstração do Algoritmo do FK

Através das estimativas inicias $\boldsymbol{x_{1^-}}$ e $\boldsymbol{P_{1^-}}$ inicia-se um processo iterativo. A simbologia adotada no texto permite escrever

$$\widehat{\boldsymbol{E}}(\boldsymbol{x_t}|\boldsymbol{Z_{t-1}}) = \boldsymbol{x_{t^-}}$$

E

$$\widehat{\boldsymbol{E}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}}) = \boldsymbol{z_{t^-}}$$

Mas utilizando a equação de medida é possível obter

$$\widehat{\boldsymbol{E}}(\boldsymbol{z_t}|\boldsymbol{x_t}) = \boldsymbol{H_t}\boldsymbol{x_t}$$

Então,

$$\boldsymbol{z_{t^-}} = \boldsymbol{H_t}\widehat{\boldsymbol{E}}(\boldsymbol{x_t}|\boldsymbol{Z_{t-1}}) = \boldsymbol{H_t}\boldsymbol{x_{t^-}}$$

Calculando o erro de estimativa para a variável observável

$$\boldsymbol{z_t} - \boldsymbol{z_{t^-}} = \boldsymbol{H_t}\boldsymbol{x_t} + \boldsymbol{u_t} - \boldsymbol{H_t}\boldsymbol{x_{t^-}} = \boldsymbol{H_t}(\boldsymbol{x_t} - \boldsymbol{x_{t^-}}) + \boldsymbol{u_t}$$

Logo, a matriz de variâncias-covariâncias da estimativa para a variável observável será

$$\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})'] = \boldsymbol{E}[(\boldsymbol{H_t}(\boldsymbol{x_t} - \boldsymbol{x_{t^-}}) + \boldsymbol{u_t})(\boldsymbol{H_t}(\boldsymbol{x_t} - \boldsymbol{x_{t^-}}) + \boldsymbol{u_t})']$$

$$\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})'] = \boldsymbol{E}[(\boldsymbol{H_t}(\boldsymbol{x_t} - \boldsymbol{x_{t^-}}) + \boldsymbol{u_t})((\boldsymbol{x_t} - \boldsymbol{x_{t^-}})'\boldsymbol{H_t}' + \boldsymbol{u_t}')]$$

$$\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})'] = \boldsymbol{E}[\boldsymbol{H_t}(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})'\boldsymbol{H_t}' + \boldsymbol{H_t}(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})\boldsymbol{u_t}' + \boldsymbol{u_t}(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})'\boldsymbol{H_t}' + \boldsymbol{u_t}\boldsymbol{u_t}']$$

No entanto, como o erro aleatório é descorrelacionado com as variáveis de estado, o valor esperado dos termos em que estes estão juntos vai para zero.

$$\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})'] = \boldsymbol{H_t}\boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})']\boldsymbol{H_t}' + \boldsymbol{E}[\boldsymbol{u_t}\boldsymbol{u_t}']$$

Ao longo do texto definiu-se que a matriz $\boldsymbol{P}$ era a matriz de variâncias-covariâncias do erro de estimativa para as variáveis não observáveis. Neste sentido, pode-se escrever que

$$\boldsymbol{P}_{(\boldsymbol{t+1})^-} = \boldsymbol{E}[(\boldsymbol{x_{t+1}} - \boldsymbol{x_{t+1^-}})(\boldsymbol{x_{t+1}} - \boldsymbol{x_{t+1^-}})']$$

Portanto,

$$\boldsymbol{P_{t^-}} = \boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})']$$

Então,

$$\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})'] = \boldsymbol{H_t}\boldsymbol{P_{t^-}}\boldsymbol{H_t}' + \boldsymbol{E}[\boldsymbol{u_t}\boldsymbol{u_t}']$$

O segundo termo do lado direito da equação acima representa exatamente a matriz de variâncias-covariâncias do erro da equação de medida. Desta forma, pode-se escrever

$$\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})'] = \boldsymbol{H_t}\boldsymbol{P_{t^-}}\boldsymbol{H_t}' + \boldsymbol{R_t}$$

Assim, obteve-se as relações que representam a estimativa para a variável observável e sua variância na data t quando conhece-se a observação de $t - 1$.

O próximo passo é obter relações para a correção das estimativas para as variáveis de estado e suas variâncias. Através da definição adotada

$$\widehat{\boldsymbol{E}}(\boldsymbol{x_t}|\boldsymbol{Z_t}) = \boldsymbol{x_{t^+}}$$

Usando a relação de atualização de uma previsão linear encontrada em Apêndice 16.3.2 é possível escrever que

$$\boldsymbol{x_{t^+}} = \boldsymbol{x_{t^-}} + \{\boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})']\} \times \{\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})']\}^{-\boldsymbol{1}} \times (\boldsymbol{z_t} - \boldsymbol{z_{t^-}})$$

Onde

$$\boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})'] = \boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{H_t}(\boldsymbol{x_t} - \boldsymbol{x_{t^-}}) + \boldsymbol{u_t})']$$

$$\boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})'] = \boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})((\boldsymbol{x_t} - \boldsymbol{x_{t^-}})'\boldsymbol{H_t}' + \boldsymbol{u_t}')]$$

$$E[(x_t - x_{t^-})(z_t - z_{t^-})'] = E[(x_t - x_{t^-})(x_t - x_{t^-})'H_t'] + E[(x_t - x_{t^-})u_t']$$

Como o erro da equação de medida é descorrelacionado com o erro de estimativa para as variáveis de estado

$$E[(x_t - x_{t^-})(z_t - z_{t^-})'] = E[(x_t - x_{t^-})(x_t - x_{t^-})']H_t'$$

Aplicando a definição de P_{t^-} acima

$$E[(x_t - x_{t^-})(z_t - z_{t^-})'] = P_{t^-}H_t'$$

Voltando à equação de x_{t^+} chega-se a

$$x_{t^+} = x_{t^-} + P_{t^-}H_t' \times \{E[(z_t - z_{t^-})(z_t - z_{t^-})']\}^{-1} \times (z_t - z_{t^-})$$

Substituindo a relação encontrada para a matriz de variâncias-covariâncias do erro de estimativa para a variável observável na equação acima

$$x_{t^+} = x_{t^-} + P_{t^-}H_t' \times \{H_t P_{t^-} H_t' + R_t\}^{-1} \times (z_t - z_{t^-})$$

E, por último, substituindo z_{t^-}

$$x_{t^+} = x_{t^-} + P_{t^-}H_t' \times \{H_t P_{t^-} H_t' + R_t\}^{-1} \times (z_t - H_t x_{t^-})$$

Na equação acima o Ganho de Kalman é definido como

$$K_t = P_{t^-}H_t'(H_t P_{t^-} H_t' + R_t)^{-1}$$

Logo,

$$x_{t^+} = x_{t^-} + K_t(z_t - H_t x_{t^-})$$

Agora, para determinar-se a matriz de variâncias para o erro de estimativa necessita-se da definição

$$P_{t^+} = E[(x_t - x_{t^+})(x_t - x_{t^+})']$$

Utilizando a relação encontrada no Apêndice 16.3.2 para determinar a matriz de variâncias associada com a previsão ajustada pode-se escrever

$$P_{t^+} = E[(x_t - x_{t^-})(x_t - x_{t^-})'] - E[(x_t - x_{t^-})(z_t - z_{t^-})'] \times \{E[(z_t - z_{t^-})(z_t - z_{t^-})']\}^{-1} \times \{E[(z_t - z_{t^-})(x_t - x_{t^-})']\}$$

Logo,

$$P_{t^+} = P_{t^-} - P_{t^-}H_t'(H_t P_{t^-} H_t' + R_t)^{-1} H_t P_{t^-}$$

Ou,

$$P_{t^+} = P_{t^-} - K_t H_t P_{t^-}$$

$$P_{t^+} = (I - K_t H_t)P_{t^-}$$

A seguir, deseja-se obter uma previsão para as variáveis de estado, tal que

$$\hat{E}(x_{t+1}|Z_t) = x_{t+1^-}$$

Aplicando à equação de estado

$$\boldsymbol{x}_{t+1^-} = \boldsymbol{\phi}\widehat{\boldsymbol{E}}(\boldsymbol{x}_t|\boldsymbol{Z}_t) + \widehat{\boldsymbol{E}}(\boldsymbol{G}\boldsymbol{v}_t|\boldsymbol{Z}_t)$$

$$\boldsymbol{x}_{t+1^-} = \boldsymbol{\phi}\boldsymbol{x}_{t^+} + \boldsymbol{0}$$

Substituindo o resultado encontrado para $\boldsymbol{x}_{t^+}$ na equação acima

$$\boldsymbol{x}_{t+1^-} = \boldsymbol{\phi}[\boldsymbol{x}_{t^-} + \boldsymbol{P}_{t^-}\boldsymbol{H}_t{}' \times (\boldsymbol{H}_t\boldsymbol{P}_{t^-}\boldsymbol{H}_t{}' + \boldsymbol{R}_t)^{-1} \times (\boldsymbol{z}_t - \boldsymbol{H}_t\boldsymbol{x}_{t^-})]$$

$$\boldsymbol{x}_{t+1^-} = \boldsymbol{\phi}\boldsymbol{x}_{t^-} + \boldsymbol{\phi}\boldsymbol{P}_{t^-}\boldsymbol{H}_t{}' \times (\boldsymbol{H}_t\boldsymbol{P}_{t^-}\boldsymbol{H}_t{}' + \boldsymbol{R}_t)^{-1} \times (\boldsymbol{z}_t - \boldsymbol{H}_t\boldsymbol{x}_{t^-})$$

Logo, substituindo $\boldsymbol{K}_t$ na equação acima

$$\boldsymbol{x}_{t+1^-} = \boldsymbol{\phi}\boldsymbol{x}_{t^-} + \boldsymbol{\phi}\boldsymbol{K}_t(\boldsymbol{z}_t - \boldsymbol{H}_t\boldsymbol{x}_{t^-})$$

Ou

$$\boldsymbol{x}_{t+1^-} = \boldsymbol{\phi}[\boldsymbol{x}_{t^-} + \boldsymbol{K}_t(\boldsymbol{z}_t - \boldsymbol{H}_t\boldsymbol{x}_{t^-})]$$

Abrindo a expressão de $\boldsymbol{K}_t$

$$\boldsymbol{x}_{t+1^-} = \boldsymbol{\phi}[\boldsymbol{x}_{t^-} + \boldsymbol{P}_{t^-}\boldsymbol{H}_t{}'(\boldsymbol{H}_t\boldsymbol{P}_{t^-}\boldsymbol{H}_t{}' + \boldsymbol{R}_t)^{-1}(\boldsymbol{z}_t - \boldsymbol{H}_t\boldsymbol{x}_{t^-})]$$

Substituindo a expressão por $\boldsymbol{x}_{t^+}$

$$\boldsymbol{x}_{t+1^-} = \boldsymbol{\phi}\boldsymbol{x}_{t^+}$$

A matriz de covariâncias será dada por

$$\boldsymbol{P}_{(t+1)^-} = \boldsymbol{E}[(\boldsymbol{x}_{t+1} - \boldsymbol{x}_{t+1^-})(\boldsymbol{x}_{t+1} - \boldsymbol{x}_{t+1^-})']$$

$$\boldsymbol{P}_{(t+1)^-} = \boldsymbol{E}[(\boldsymbol{\phi}\boldsymbol{x}_t + \boldsymbol{G}\boldsymbol{v}_{t+1} - \boldsymbol{\phi}\boldsymbol{x}_{t^+})(\boldsymbol{\phi}\boldsymbol{x}_t + \boldsymbol{G}\boldsymbol{v}_{t+1} - \boldsymbol{\phi}\boldsymbol{x}_{t^+})']$$

$$\boldsymbol{P}_{(t+1)^-} = \boldsymbol{E}[(\boldsymbol{\phi}\boldsymbol{x}_t + \boldsymbol{G}\boldsymbol{v}_{t+1} - \boldsymbol{\phi}\boldsymbol{x}_{t^+})(\boldsymbol{x}_t{}'\boldsymbol{\phi}' + \boldsymbol{v}_{t+1}{}'\boldsymbol{G}' - \boldsymbol{x}_{t^+}{}'\boldsymbol{\phi}')]$$

$$\begin{aligned}\boldsymbol{P}_{(t+1)^-} = \boldsymbol{E}[&\boldsymbol{\phi}\boldsymbol{x}_t\boldsymbol{x}_t{}'\boldsymbol{\phi}' + \boldsymbol{\phi}\boldsymbol{x}_t\boldsymbol{v}_{t+1}{}'\boldsymbol{G}' - \boldsymbol{\phi}\boldsymbol{x}_t\boldsymbol{x}_{t^+}{}'\boldsymbol{\phi}' + \boldsymbol{G}\boldsymbol{v}_{t+1}\boldsymbol{x}_t{}'\boldsymbol{\phi}' +\\ &\boldsymbol{G}\boldsymbol{v}_{t+1}\boldsymbol{v}_{t+1}{}'\boldsymbol{G}' - \boldsymbol{G}\boldsymbol{v}_{t+1}\boldsymbol{x}_{t^+}{}'\boldsymbol{\phi}' - \boldsymbol{\phi}\boldsymbol{x}_{t^+}\boldsymbol{x}_t{}'\boldsymbol{\phi}' - \boldsymbol{\phi}\boldsymbol{x}_{t^+}\boldsymbol{v}_{t+1}{}'\boldsymbol{G}' +\\ &\boldsymbol{\phi}\boldsymbol{x}_{t^+}\boldsymbol{x}_{t^+}{}'\boldsymbol{\phi}']\end{aligned}$$

Como o erro aleatório da equação de estado em $t + 1$ não é correlacionado com as variáveis de estado em t, o valor esperado desses termos vai para zero.

$$\boldsymbol{P}_{(t+1)^-} = \boldsymbol{\phi}\boldsymbol{E}[(\boldsymbol{x}_t - \boldsymbol{x}_{t^+})(\boldsymbol{x}_t - \boldsymbol{x}_{t^+})']\boldsymbol{\phi}' + \boldsymbol{G}\boldsymbol{E}[\boldsymbol{v}_{t+1}\boldsymbol{v}_{t+1}{}']\boldsymbol{G}'$$

Utilizando definições anteriores

$$\boldsymbol{P}_{(t+1)^-} = \boldsymbol{\phi}\boldsymbol{P}_{t^+}\boldsymbol{\phi}' + \boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}'$$

Substituindo na equação acima a expressão encontrada para $\boldsymbol{P}_{t^+}$

$$\boldsymbol{P}_{(t+1)^-} = \boldsymbol{\phi}[\boldsymbol{P}_{t^-} - \boldsymbol{P}_{t^-}\boldsymbol{H}_t{}'(\boldsymbol{H}_t\boldsymbol{P}_{t^-}\boldsymbol{H}_t{}' + \boldsymbol{R}_t)^{-1}\boldsymbol{H}_t\boldsymbol{P}_{t^-}]\boldsymbol{\phi}' + \boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}'$$

Desta forma, demonstraram-se todas as equações de atualização.

16.3. Atualização do Vetor de Estado Após a Observação

Dado

$\boldsymbol{Y} = (Y_1, Y_2, \dots, Y_n)'$

$\boldsymbol{\Omega} = \boldsymbol{E}(\boldsymbol{YY'})$

Onde $\boldsymbol{\Omega}$ representa a matriz de covariâncias de $\boldsymbol{Y}$.

Seja,

$\boldsymbol{\bar{Y}} = \boldsymbol{A^{-1}Y}$

Então,

$\boldsymbol{E}(\boldsymbol{\bar{Y}\bar{Y}'}) = \boldsymbol{E}(\boldsymbol{A^{-1}YY'}[\boldsymbol{A'}]^{-\mathbf{1}}) = \boldsymbol{A^{-1}E}(\boldsymbol{YY'})[\boldsymbol{A'}]^{-\mathbf{1}} = \boldsymbol{A^{-1}\Omega}[\boldsymbol{A'}]^{-\mathbf{1}}$

No entanto, no Apêndice 16.3.1 prova-se que

$\boldsymbol{\Omega} = \boldsymbol{ADA'}$

Logo,

$\boldsymbol{E}(\boldsymbol{\bar{Y}\bar{Y}'}) = \boldsymbol{A^{-1}ADA'}[\boldsymbol{A'}]^{-\mathbf{1}} = \boldsymbol{D}$

Onde

$E\left(\bar{Y}_i\bar{Y}_j\right) = \begin{cases} d_{ii} & para\ i = j \\ 0 & para\ i \neq j \end{cases}$

Pré-multiplicando a equação $\boldsymbol{\bar{Y}} = \boldsymbol{A^{-1}Y}$ por $\boldsymbol{A}$

$\boldsymbol{A\bar{Y}} = \boldsymbol{Y}$

Também do Apêndice 16.3.1 vem a definição da matriz $\boldsymbol{A}$, o que permite escrever

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \Omega_{21}{\Omega_{11}}^{-1} & 1 & 0 & \dots & 0 \\ \Omega_{31}{\Omega_{11}}^{-1} & h_{32}{h_{22}}^{-1} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Omega_{n1}{\Omega_{11}}^{-1} & h_{n2}{h_{11}}^{-1} & k_{n3}{k_{33}}^{-1} & \dots & 1 \end{bmatrix} \begin{bmatrix} \bar{Y}_1 \\ \bar{Y}_2 \\ \bar{Y}_3 \\ \vdots \\ \bar{Y}_n \end{bmatrix} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix}$$

Da equação matricial acima é possível escrever

$\bar{Y}_1 = Y_1$

E também

$\Omega_{21}{\Omega_{11}}^{-1}\bar{Y}_1 + \bar{Y}_2 = Y_2$

Substituindo a equação de $\bar{Y}_1$ na equação acima

$\Omega_{21}{\Omega_{11}}^{-1}Y_1 + \bar{Y}_2 = Y_2$

$\bar{Y}_2 = Y_2 - \Omega_{21}{\Omega_{11}}^{-1}Y_1$

Como,

$E(\bar{Y}_2\bar{Y}_1) = 0$

Então

$E[(Y_2 - \alpha Y_1)Y_1] = 0$

Onde

$\alpha = \Omega_{21}{\Omega_{11}}^{-1}$

A variável α representa o coeficiente linear da projeção de Y_2 sobre Y_1. De modo geral o valor da linha i da coluna 1 de $\boldsymbol{A}$, $\Omega_{i1}{\Omega_{11}}^{-1}$, representa o coeficiente da projeção linear de Y_i sobre Y_1. Como $\bar{Y}_2$ é o resíduo da projeção de Y_2 em Y_1, d_{22} dará o erro médio quadrático da projeção.

$E\left({\bar{Y}_2}^2\right) = d_{22} = \Omega_{22} - \Omega_{21}{\Omega_{11}}^{-1}\Omega_{12}$

A relação encontrada para d_{22} vem do Apêndice 16.3.1.

Voltando à equação matricial $\boldsymbol{A\bar{Y}} = \boldsymbol{Y}$, pode-se escrever

$\Omega_{31}{\Omega_{11}}^{-1}\bar{Y}_1 + h_{32}{h_{22}}^{-1}\bar{Y}_2 + \bar{Y}_3 = Y_3$

Substituindo as relações encontradas anteriormente para $\bar{Y}_1$ e $\bar{Y}_2$ na equação acima chega-se a

$\Omega_{31}{\Omega_{11}}^{-1}Y_1 + h_{32}{h_{22}}^{-1}\left(Y_2 - \Omega_{21}{\Omega_{11}}^{-1}Y_1\right) + \bar{Y}_3 = Y_3$

$\bar{Y}_3 = Y_3 - \Omega_{31}{\Omega_{11}}^{-1}Y_1 - h_{32}{h_{22}}^{-1}\left(Y_2 - \Omega_{21}{\Omega_{11}}^{-1}Y_1\right)$

A covariância entre $\bar{Y}_3$ e $\bar{Y}_1$ e $\bar{Y}_2$ será dada por

$E\left[\left(Y_3 - \Omega_{31}{\Omega_{11}}^{-1}Y_1 - h_{32}{h_{22}}^{-1}\left(Y_2 - \Omega_{21}{\Omega_{11}}^{-1}Y_1\right)\right)\bar{Y}_j\right] = 0$

Para $j = 1, 2$

Pois a matriz $\boldsymbol{D}$ é diagonal.

Como $\bar{Y}_3$ é o resíduo da projeção linear de Y_3 em função de Y_2 e Y_1, então é possível escrever a previsão para Y_3 como

$\hat{P}(Y_3|Y_2, Y_1) = \Omega_{31}{\Omega_{11}}^{-1}Y_1 + h_{32}{h_{22}}^{-1}\left(Y_2 - \Omega_{21}{\Omega_{11}}^{-1}Y_1\right)$

E o erro médio quadrático da previsão será dado pela variância de $\bar{Y}_3$

$E\left[\left(Y_3 - \hat{P}(Y_3|Y_2, Y_1)\right)^2\right] = d_{33} = h_{33} - h_{32}{h_{22}}^{-1}h_{23}$

Onde d_{33} é obtido no Apêndice 16.3.1.

Se apenas a informação Y_1 for conhecida a previsão para Y_3 será

$\hat{P}(Y_3|Y_1) = \Omega_{31}{\Omega_{11}}^{-1}Y_1$

Da mesma forma a previsão para Y_2 considerando conhecido apenas Y_1 será

$\hat{P}(Y_2|Y_1) = \Omega_{21}{\Omega_{11}}^{-1}Y_1$

Desta forma, é possível reescrever a previsão para Y_3 baseada em Y_2 e Y_1, tal que

$\hat{P}(Y_3|Y_2, Y_1) = \hat{P}(Y_3|Y_1) + h_{32}{h_{22}}^{-1}\left(Y_2 - \hat{P}(Y_2|Y_1)\right)$

A equação acima explicita que a previsão para Y_3, baseada em Y_2 e Y_1, parte de uma estimativa inicial para Y_3 baseada em Y_1, e acrescenta um componente não previsto multiplicado por um coeficiente. Para compreender a natureza deste coeficiente define-se

$$\overline{\boldsymbol{Y}}(\boldsymbol{1}) = \boldsymbol{E_1 Y}$$

Onde $\overline{\boldsymbol{Y}}(\boldsymbol{1})$ possui dimensão $n \times 1$ e $\boldsymbol{E_1}$ será descrita mais adiante quando for feita a fatoração triangular de uma matriz simétrica positiva definida. A matriz de covariâncias de $\overline{\boldsymbol{Y}}(\boldsymbol{1})$ pode ser definida por

$$\boldsymbol{E}\{\overline{\boldsymbol{Y}}(\boldsymbol{1})[\overline{\boldsymbol{Y}}(\boldsymbol{1})]'\} = \boldsymbol{E}\{\boldsymbol{E_1 Y}\,\boldsymbol{Y}'\boldsymbol{E_1}\,'\} = \boldsymbol{E_1 \Omega E_1}\,'$$

Mas do Apêndice 16.3.1 vem

$$\boldsymbol{H} = \boldsymbol{E_1 \Omega E_1}\,'$$

Além disso, multiplicando $\boldsymbol{E_1}$ por $\boldsymbol{Y}$, $\overline{\boldsymbol{Y}}(\boldsymbol{1})$ será dado por

$$\overline{\boldsymbol{Y}}(\boldsymbol{1}) = \begin{bmatrix} Y_1 \\ Y_2 - \Omega_{21}\Omega_{11}{}^{-1}Y_1 \\ Y_3 - \Omega_{31}\Omega_{11}{}^{-1}Y_1 \\ \vdots \\ Y_n - \Omega_{n1}\Omega_{11}{}^{-1}Y_1 \end{bmatrix}$$

Cada linha do vetor acima representa o resíduo da projeção de Y_i em função de Y_1. Desta forma, como $\boldsymbol{H}$ foi definida como a matriz de covariâncias de $\overline{\boldsymbol{Y}}(\boldsymbol{1})$, e $\overline{\boldsymbol{Y}}(\boldsymbol{1})$ representa um vetor de resíduos, $\boldsymbol{H}$ será a matriz de covariâncias dos resíduos da projeção de cada uma das variáveis Y_i em função de Y_1. Neste sentido,

$$h_{22} = E\left[\left(Y_2 - \hat{P}(Y_2|Y_1)\right)^2\right]$$

E

$$h_{32} = E\left\{\left[Y_3 - \hat{P}(Y_3|Y_1)\right]\left[Y_2 - \hat{P}(Y_2|Y_1)\right]\right\}$$

Utilizando as definições acima pode-se escrever a previsão de Y_3 em função de Y_2 e Y_1 como

$$\hat{P}(Y_3|Y_2, Y_1) = \hat{P}(Y_3|Y_1) + E\left\{\left[Y_3 - \hat{P}(Y_2|Y_1)\right]\left[Y_2 - \hat{P}(Y_2|Y_1)\right]\right\} \times \left\{\left(Y_2 - \hat{P}(Y_2|Y_1)\right)^2\right\}^{-1} \times \left(Y_2 - \hat{P}(Y_2|Y_1)\right)$$

Em geral, para $i > 2$ o coeficiente de Y_2 em uma projeção de Y_i em Y_2 e Y_1 é dado pelo i-ésimo elemento da segunda coluna da matriz A. Para qualquer $i > j$ os coeficientes de Y_j em uma projeção de Y_i em Y_j, Y_{j-1}, ..., Y_1, são dados pelo elemento da linha i coluna j da matriz A. O elemento d_{ii} será o erro médio quadrático para a projeção de Y_i em Y_i, Y_{i-1}, ..., Y_1.

16.3.1. Fatoração Triangular de Uma Matriz Simétrica Positiva Definida

As equações utilizadas no processo de atualização descrito acima são definidas por um procedimento conhecido como fatoração triangular de uma matriz simétrica positiva definida. A seguir define-se tal procedimento para que seja conhecida a forma de gerar as matrizes do processo de atualização.

Seja,

$$\boldsymbol{\Omega} = \boldsymbol{ADA'}$$

Onde:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{21} & 1 & 0 & \dots & 0 \\ a_{31} & a_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & 1 \end{bmatrix}$$

$$\boldsymbol{D} = \begin{bmatrix} d_{11} & 0 & 0 & \dots & 0 \\ 0 & d_{22} & 0 & \dots & 0 \\ 0 & 0 & d_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_{nn} \end{bmatrix}$$

$$\boldsymbol{\Omega} = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} & \dots & \Omega_{1n} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} & \dots & \Omega_{2n} \\ \Omega_{31} & \Omega_{32} & \Omega_{33} & \cdots & \Omega_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Omega_{n1} & \Omega_{n2} & \Omega_{n3} & \cdots & \Omega_{nn} \end{bmatrix}$$

A matriz $\boldsymbol{D}$ apresenta os elementos da diagonal principal tais que $d_{ii} > 0$. No caso da matriz $\boldsymbol{\Omega}$, ela se caracteriza por uma matriz positiva definida, e portanto, $\boldsymbol{x'\Omega x} > \boldsymbol{0}$, onde $\boldsymbol{x}$ apresenta dimensão $n \times 1$. Por ser uma matriz simétrica $\Omega_{ij} = \Omega_{ji}$.

Com o objetivo de determinar as matrizes $\boldsymbol{A}$ e $\boldsymbol{D}$, a matriz $\boldsymbol{\Omega}$ será multiplicada de tal forma a se tornar uma matriz diagonal. Nesse sentido, primeiramente com o objetivo de zerar a primeira coluna abaixo da diagonal principal a matriz $\boldsymbol{\Omega}$ será multiplicada por $\boldsymbol{E_1}$, tal que

$$\boldsymbol{E_1} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\Omega_{21}{\Omega_{11}}^{-1} & 1 & 0 & \dots & 0 \\ -\Omega_{31}{\Omega_{11}}^{-1} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\Omega_{n1}{\Omega_{11}}^{-1} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Com o objetivo de tornar a primeira linha igual a zero após a diagonal principal multiplica-se a $\boldsymbol{E_1\Omega}$ por $\boldsymbol{E_1}'$. Assim obtém-se

$$\boldsymbol{E_1\Omega E_1}' = \boldsymbol{H}$$

Tal que,

$$\boldsymbol{H} = \begin{bmatrix} h_{11} & 0 & 0 & \dots & 0 \\ 0 & h_{22} & h_{23} & \dots & h_{2n} \\ 0 & h_{32} & h_{33} & \dots & h_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & h_{n2} & h_{n3} & \cdots & h_{nn} \end{bmatrix}$$

$$\boldsymbol{H} =$$

$$\begin{bmatrix} \Omega_{11} & 0 & 0 & \dots & 0 \\ 0 & \Omega_{22}-\Omega_{21}\Omega_{11}{}^{-1}\Omega_{12} & \Omega_{23}-\Omega_{21}\Omega_{11}{}^{-1}\Omega_{13} & \dots & \Omega_{2n}-\Omega_{21}\Omega_{11}{}^{-1}\Omega_{1n} \\ 0 & \Omega_{32}-\Omega_{31}\Omega_{11}{}^{-1}\Omega_{12} & \Omega_{33}-\Omega_{31}\Omega_{11}{}^{-1}\Omega_{13} & \cdots & \Omega_{3n}-\Omega_{31}\Omega_{11}{}^{-1}\Omega_{1n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \Omega_{n2}-\Omega_{n1}\Omega_{11}{}^{-1}\Omega_{12} & \Omega_{n3}-\Omega_{n1}\Omega_{11}{}^{-1}\Omega_{13} & \cdots & \Omega_{nn}-\Omega_{n1}\Omega_{11}{}^{-1}\Omega_{1n} \end{bmatrix}$$

O mesmo procedimento deve ser efetuado para zerar a segunda coluna abaixo da diagonal principal e a segunda linha após a diagonal principal. Assim chega-se a

$$\boldsymbol{E_2 H E_2}' = \boldsymbol{K}$$

Onde

$$\boldsymbol{E_2} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & -h_{32}h_{22}{}^{-1} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -h_{n2}h_{22}{}^{-1} & 0 & \cdots & 1 \end{bmatrix}$$

$$\boldsymbol{K} = \begin{bmatrix} h_{11} & 0 & 0 & \dots & 0 \\ 0 & h_{22} & 0 & \dots & 0 \\ 0 & 0 & h_{33}-h_{32}h_{22}{}^{-1}h_{23} & \dots & h_{3n}-h_{32}h_{22}{}^{-1}h_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & h_{n3}-h_{n2}h_{22}{}^{-1}h_{23} & \cdots & h_{nn}-h_{n2}h_{22}{}^{-1}h_{2n} \end{bmatrix}$$

Utilizando a mesma técnica para zerar as outras colunas e linhas abaixo e após a diagonal principal respectivamente chega-se a seguinte matriz diagonal, a qual será igual à matriz $\boldsymbol{D}$, por construção. Pode-se escrever

$$\boldsymbol{E_{n-1}} \dots \boldsymbol{E_2 E_1 \Omega E_1}'\boldsymbol{E_2}' \dots \boldsymbol{E_{n-1}}' = \boldsymbol{D}$$

Onde

$$D = \begin{bmatrix} \Omega_{11} & 0 & 0 & \dots & 0 \\ 0 & \Omega_{22} - \Omega_{21}{\Omega_{11}}^{-1}\Omega_{12} & 0 & \dots & 0 \\ 0 & 0 & h_{33} - h_{32}{h_{22}}^{-1}h_{23} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & c_{nn} - c_{n,n-1}{c_{n-1,n-1}}^{-1}c_{n-1,n} \end{bmatrix}$$

Como $\boldsymbol{E_j}$ é uma matriz com uns na diagonal, valores diferentes de zero na coluna $\boldsymbol{j}$ abaixo da diagonal principal e zeros nas outras posições, $\boldsymbol{E_j}^{-1}$ existe. Nomeando,

$$\boldsymbol{A} = (\boldsymbol{E_{n-1}} \dots \boldsymbol{E_2 E_1})^{-1} = \boldsymbol{E_1}^{-1}\boldsymbol{E_2}^{-1} \dots \boldsymbol{E_{n-1}}^{-1}$$

Logo, a matriz $\boldsymbol{A}$ será

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\Omega_{21}{\Omega_{11}}^{-1} & 1 & 0 & \dots & 0 \\ -\Omega_{31}{\Omega_{11}}^{-1} & -h_{32}{h_{22}}^{-1} & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\Omega_{n1}{\Omega_{11}}^{-1} & -h_{n2}{h_{22}}^{-1} & -k_{n3}{k_{33}}^{-1} & \dots & 1 \end{bmatrix}$$

Portanto, voltando a uma equação anterior e pré multiplicando-a por $\boldsymbol{A}$ e pós multiplicando-a por $\boldsymbol{A'}$

$$\boldsymbol{A}(\boldsymbol{E_{n-1}} \dots \boldsymbol{E_2 E_1})\boldsymbol{\Omega}(\boldsymbol{E_1}'\boldsymbol{E_2}' \dots \boldsymbol{E_{n-1}}')\boldsymbol{A}' = \boldsymbol{ADA}'$$

$$(\boldsymbol{E_{n-1}} \dots \boldsymbol{E_2 E_1})^{-1}(\boldsymbol{E_{n-1}} \dots \boldsymbol{E_2 E_1})\boldsymbol{\Omega}(\boldsymbol{E_1}'\boldsymbol{E_2}' \dots \boldsymbol{E_{n-1}}')(\boldsymbol{E_1}'\boldsymbol{E_2}' \dots \boldsymbol{E_{n-1}}')^{-1} = \boldsymbol{ADA}'$$

Logo,

$$\boldsymbol{\Omega} = \boldsymbol{ADA}'$$

Desta forma, é possível conhecer as matrizes $\boldsymbol{A}$ e $\boldsymbol{D}$ que foram utilizadas na obtenção da equação de atualização da previsão linear.

16.3.2.
Fatoração Triangular em Bloco

Para o caso deste estudo, em que há mais de uma série de variáveis, a matriz de covariâncias é dada, por exemplo, para o caso de duas séries pela relação

$$\boldsymbol{\Omega} = \begin{bmatrix} \boldsymbol{E}(\boldsymbol{Y_1 Y_1}') & \boldsymbol{E}(\boldsymbol{Y_1 Y_2}') \\ \boldsymbol{E}(\boldsymbol{Y_2 Y_1}') & \boldsymbol{E}(\boldsymbol{Y_2 Y_2}') \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Omega_{11}} & \boldsymbol{\Omega_{12}} \\ \boldsymbol{\Omega_{21}} & \boldsymbol{\Omega_{22}} \end{bmatrix}$$

As dimensões dos termos da matriz acima são $n_1 \times n_1$ para $\boldsymbol{\Omega_{11}}$, $n_2 \times n_2$ para $\boldsymbol{\Omega_{22}}$, $n_1 \times n_2$ para $\boldsymbol{\Omega_{12}}$ e $n_2 \times n_1$ para $\boldsymbol{\Omega_{21}}$.

A matriz que irá igualar a zero o termo $\boldsymbol{\Omega_{21}}$ quando pré multiplicar $\boldsymbol{\Omega}$ será

$$\bar{\boldsymbol{E}}_{\boldsymbol{1}} = \begin{bmatrix} \boldsymbol{I}_{\boldsymbol{n_1}} & \boldsymbol{0} \\ -\boldsymbol{\Omega}_{\boldsymbol{21}}\boldsymbol{\Omega}_{\boldsymbol{11}}{}^{-1} & \boldsymbol{I}_{\boldsymbol{n_2}} \end{bmatrix}$$

E quando pós multiplicar por $\bar{\boldsymbol{E}}_{\boldsymbol{1}}{}'$ o resultado será

$$\bar{\boldsymbol{E}}_{\boldsymbol{1}}\boldsymbol{\Omega}\bar{\boldsymbol{E}}_{\boldsymbol{1}}{}' = \begin{bmatrix} \boldsymbol{I}_{\boldsymbol{n_1}} & \boldsymbol{0} \\ -\boldsymbol{\Omega}_{\boldsymbol{21}}\boldsymbol{\Omega}_{\boldsymbol{11}}{}^{-1} & \boldsymbol{I}_{\boldsymbol{n_2}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{\Omega}_{\boldsymbol{11}} & \boldsymbol{\Omega}_{\boldsymbol{12}} \\ \boldsymbol{\Omega}_{\boldsymbol{21}} & \boldsymbol{\Omega}_{\boldsymbol{22}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{I}_{\boldsymbol{n_1}} & -\boldsymbol{\Omega}_{\boldsymbol{21}}\boldsymbol{\Omega}_{\boldsymbol{11}}{}^{-1} \\ \boldsymbol{0} & \boldsymbol{I}_{\boldsymbol{n_2}} \end{bmatrix}$$

$$\bar{\boldsymbol{E}}_{\boldsymbol{1}}\boldsymbol{\Omega}\bar{\boldsymbol{E}}_{\boldsymbol{1}}{}' = \begin{bmatrix} \boldsymbol{\Omega}_{\boldsymbol{11}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Omega}_{\boldsymbol{22}} - \boldsymbol{\Omega}_{\boldsymbol{21}}\boldsymbol{\Omega}_{\boldsymbol{11}}{}^{-1}\boldsymbol{\Omega}_{\boldsymbol{12}} \end{bmatrix}$$

É possível chamar a relação acima de $\bar{\boldsymbol{D}}$, que é uma matriz em bloco diagonal. Logo,

$$\bar{\boldsymbol{D}} = \bar{\boldsymbol{E}}_{\boldsymbol{1}}\boldsymbol{\Omega}\bar{\boldsymbol{E}}_{\boldsymbol{1}}{}'$$

Definindo

$$\bar{\boldsymbol{A}} \equiv \bar{\boldsymbol{E}}_{\boldsymbol{1}}{}^{-\boldsymbol{1}} = \begin{bmatrix} \boldsymbol{I}_{\boldsymbol{n_1}} & \boldsymbol{0} \\ \boldsymbol{\Omega}_{\boldsymbol{21}}\boldsymbol{\Omega}_{\boldsymbol{11}}{}^{-1} & \boldsymbol{I}_{\boldsymbol{n_2}} \end{bmatrix}$$

Se a relação $\bar{\boldsymbol{E}}_{\boldsymbol{1}}\boldsymbol{\Omega}\bar{\boldsymbol{E}}_{\boldsymbol{1}}{}'$ for pré multiplicada por $\bar{\boldsymbol{A}}$ e pós multiplicada por $\bar{\boldsymbol{A}}'$

$$\bar{\boldsymbol{A}}\bar{\boldsymbol{E}}_{\boldsymbol{1}}\boldsymbol{\Omega}\bar{\boldsymbol{E}}_{\boldsymbol{1}}{}'\bar{\boldsymbol{A}}' = \boldsymbol{\Omega}$$

Ou,

$$\boldsymbol{\Omega} = \bar{\boldsymbol{A}}\bar{\boldsymbol{D}}\bar{\boldsymbol{A}}'$$

Esse procedimento é semelhante ao aplicado anteriormente na fatoração triangular, a única diferença é que ao invés de $\bar{\boldsymbol{D}}$ ser uma matriz diagonal, ela é uma matriz em bloco diagonal.

A matriz $\bar{\boldsymbol{D}}$ será a matriz de covariâncias de $\bar{\boldsymbol{Y}}$, o que implica que a covariância entre qualquer elemento de $\bar{\boldsymbol{Y}}_{\boldsymbol{2}}$ e $\bar{\boldsymbol{Y}}_{\boldsymbol{1}}$ é igual a zero, e onde $\bar{\boldsymbol{Y}}$ é dado por

$$\bar{\boldsymbol{Y}} = \bar{\boldsymbol{A}}^{-\boldsymbol{1}}\boldsymbol{Y} = \begin{bmatrix} \bar{\boldsymbol{Y}}_{\boldsymbol{1}} \\ \bar{\boldsymbol{Y}}_{\boldsymbol{2}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_{\boldsymbol{n_1}} & \boldsymbol{0} \\ -\boldsymbol{\Omega}_{\boldsymbol{21}}\boldsymbol{\Omega}_{\boldsymbol{11}}{}^{-1} & \boldsymbol{I}_{\boldsymbol{n_2}} \end{bmatrix} \begin{bmatrix} \boldsymbol{Y}_{\boldsymbol{1}} \\ \boldsymbol{Y}_{\boldsymbol{2}} \end{bmatrix}$$

Assim,

$$\bar{\boldsymbol{Y}}_{\boldsymbol{1}} = \boldsymbol{Y}_{\boldsymbol{1}}$$

$$\bar{\boldsymbol{Y}}_{\boldsymbol{2}} = \boldsymbol{Y}_{\boldsymbol{2}} - \boldsymbol{\Omega}_{\boldsymbol{21}}\boldsymbol{\Omega}_{\boldsymbol{11}}{}^{-\boldsymbol{1}}\boldsymbol{Y}_{\boldsymbol{1}}$$

Pode-se escrever portanto que

$$\widehat{\boldsymbol{P}}(\boldsymbol{Y}_{\boldsymbol{2}}|\boldsymbol{Y}_{\boldsymbol{1}}) = \boldsymbol{\Omega}_{\boldsymbol{21}}\boldsymbol{\Omega}_{\boldsymbol{11}}{}^{-\boldsymbol{1}}\boldsymbol{Y}_{\boldsymbol{1}}$$

O erro médio quadrático, em inglês *Mean Squared Error (MSE)*, da projeção linear será dado por

$$\boldsymbol{E}\left\{\left[\boldsymbol{Y}_{\boldsymbol{2}} - \widehat{\boldsymbol{P}}(\boldsymbol{Y}_{\boldsymbol{2}}|\boldsymbol{Y}_{\boldsymbol{1}})\right]\left[\boldsymbol{Y}_{\boldsymbol{2}} - \widehat{\boldsymbol{P}}(\boldsymbol{Y}_{\boldsymbol{2}}|\boldsymbol{Y}_{\boldsymbol{1}})\right]'\right\} = \boldsymbol{E}\left(\bar{\boldsymbol{Y}}_{\boldsymbol{2}}\bar{\boldsymbol{Y}}_{\boldsymbol{2}}{}'\right) = \bar{\boldsymbol{D}}_{\boldsymbol{22}}$$

Logo,

$$\boldsymbol{E}\left\{\left[\boldsymbol{Y_2}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_2}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]'\right\}=\boldsymbol{\Omega_{22}}-\boldsymbol{\Omega_{21}}\boldsymbol{\Omega_{11}}^{-1}\boldsymbol{\Omega_{12}}$$

Para o caso em que são consideradas três séries $\boldsymbol{Y_1}$, $\boldsymbol{Y_2}$ e $\boldsymbol{Y_3}$, cada uma com dimensões $n_1 \times 1$, $n_2 \times 1$ e $n_3 \times 1$ respectivamente, a fatoração triangular será dada por

$$\begin{bmatrix}\boldsymbol{\Omega_{11}} & \boldsymbol{\Omega_{12}} & \boldsymbol{\Omega_{13}}\\ \boldsymbol{\Omega_{21}} & \boldsymbol{\Omega_{22}} & \boldsymbol{\Omega_{23}}\\ \boldsymbol{\Omega_{31}} & \boldsymbol{\Omega_{32}} & \boldsymbol{\Omega_{33}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{I_{n_1}} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{\Omega_{21}}\boldsymbol{\Omega_{11}}^{-1} & \boldsymbol{I_{n_2}} & \boldsymbol{0}\\ \boldsymbol{\Omega_{31}}\boldsymbol{\Omega_{11}}^{-1} & \mathbf{H_{32}}\mathbf{H_{22}}^{-1} & \boldsymbol{I_{n_3}}\end{bmatrix}\times$$

$$\begin{bmatrix}\boldsymbol{\Omega_{11}} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \mathbf{H_{22}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \mathbf{H_{33}}-\mathbf{H_{32}}\mathbf{H_{22}}^{-1}\mathbf{H_{23}}\end{bmatrix}\begin{bmatrix}\boldsymbol{I_{n_1}} & \boldsymbol{\Omega_{11}}^{-1}\boldsymbol{\Omega_{12}} & \boldsymbol{\Omega_{11}}^{-1}\boldsymbol{\Omega_{13}}\\ \boldsymbol{0} & \boldsymbol{I_{n_2}} & \mathbf{H_{22}}^{-1}\boldsymbol{\Omega_{23}}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I_{n_3}}\end{bmatrix}$$

Onde

$$\mathbf{H_{22}}=\boldsymbol{\Omega_{22}}-\boldsymbol{\Omega_{21}}\boldsymbol{\Omega_{11}}^{-1}\boldsymbol{\Omega_{12}}$$

$$\mathbf{H_{33}}=\boldsymbol{\Omega_{33}}-\boldsymbol{\Omega_{31}}\boldsymbol{\Omega_{11}}^{-1}\boldsymbol{\Omega_{13}}$$

$$\mathbf{H_{23}}=\mathbf{H_{32}}'=\boldsymbol{\Omega_{23}}-\boldsymbol{\Omega_{21}}\boldsymbol{\Omega_{11}}^{-1}\boldsymbol{\Omega_{13}}$$

O caso em que há três séries de variáveis é o qual se utiliza na atualização da variável de estado no caso do FK.

Neste caso, a equação de projeção linear de $\boldsymbol{Y_3}$ em função de $\boldsymbol{Y_1}$ e $\boldsymbol{Y_2}$ será dada por

$$\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_2},\boldsymbol{Y_1})=\boldsymbol{\Omega_{31}}\boldsymbol{\Omega_{11}}^{-1}\boldsymbol{Y_1}+\boldsymbol{H_{32}}\boldsymbol{H_{22}}^{-1}\left(\boldsymbol{Y_2}-\boldsymbol{\Omega_{21}}\boldsymbol{\Omega_{11}}^{-1}\boldsymbol{Y_1}\right)$$

$$\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_2},\boldsymbol{Y_1})=\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_1})+\boldsymbol{H_{32}}\boldsymbol{H_{22}}^{-1}\left(\boldsymbol{Y_2}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right)$$

Onde

$$\boldsymbol{H_{22}}=\boldsymbol{E}\left\{\left[\boldsymbol{Y_2}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_2}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]'\right\}$$

$$\boldsymbol{H_{32}}=\boldsymbol{E}\left\{\left[\boldsymbol{Y_3}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_2}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]'\right\}$$

A matriz do MSE representa a matriz de covariância do erro de projeção $\boldsymbol{\overline{Y}_3}$ que é dada por

$$\boldsymbol{E}\left\{\left[\boldsymbol{Y_3}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_2},\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_3}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_2},\boldsymbol{Y_1})\right]'\right\}=\boldsymbol{\overline{D}_{33}}$$

$$\boldsymbol{E}\left\{\left[\boldsymbol{Y_3}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_2},\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_3}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_2},\boldsymbol{Y_1})\right]'\right\}=\mathbf{H_{33}}-\mathbf{H_{32}}\mathbf{H_{22}}^{-1}\mathbf{H_{23}}$$

Onde

$$\mathbf{H_{33}}=\boldsymbol{E}\left\{\left[\boldsymbol{Y_3}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_3}-\boldsymbol{\widehat{P}}(\boldsymbol{Y_3}|\,\boldsymbol{Y_1})\right]'\right\}$$

As equações encontradas acima para a atualização da estimativa serão utilizadas no algoritmo do FK utilizando a seguinte equivalência de variáveis.

$\boldsymbol{x_t} \equiv \boldsymbol{Y_3}$

$\boldsymbol{z_t} \equiv \boldsymbol{Y_2}$

$\boldsymbol{Z_{t-1}} \equiv \boldsymbol{Y_1}$

Onde $\boldsymbol{Z_{t-1}}$ significa que toda a série de $\boldsymbol{z_t}$ até $t-1$ é conhecida.

Desta forma, a equação de projeção que pode ser escrita como

$$\begin{aligned}\widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_2},\boldsymbol{Y_1}) = \widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_1}) + \boldsymbol{E}\left\{\left[\boldsymbol{Y_3} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_2} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]'\right\} \\ \times\left(\boldsymbol{E}\left\{\left[\boldsymbol{Y_2} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_2} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]'\right\}\right)^{-1} \\ \times\left(\boldsymbol{Y_2} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right)\end{aligned}$$

Poderá ser reescrita como

$$\begin{aligned}\widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{z_t},\boldsymbol{Z_{t-1}}) = \widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{Z_{t-1}}) + \boldsymbol{E}\left\{\left[\boldsymbol{x_t} - \widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{Z_{t-1}})\right]\left[\boldsymbol{z_t} - \widehat{\boldsymbol{P}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}})\right]'\right\} \\ \times\left(\boldsymbol{E}\left\{\left[\boldsymbol{z_t} - \widehat{\boldsymbol{P}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}})\right]\left[\boldsymbol{z_t} - \widehat{\boldsymbol{P}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}})\right]'\right\}\right)^{-1} \\ \times\left(\boldsymbol{z_t} - \widehat{\boldsymbol{P}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}})\right)\end{aligned}$$

No entanto, escrevendo com a nomenclatura adotada nesta dissertação a equação acima seria

$$\begin{aligned}\boldsymbol{x_{t^+}} = \boldsymbol{x_{t^-}} + \{\boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})']\} \times \{\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})']\}^{-1} \\ \times (\boldsymbol{z_t} - \boldsymbol{z_{t^-}})\end{aligned}$$

Pois,

$\widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{z_t},\boldsymbol{Z_{t-1}})$ representa a projeção linear para $\boldsymbol{x_t}$ após a observação de $\boldsymbol{z_t}$, o que se chamou de $\boldsymbol{x_{t^+}}$.

$\widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{Z_{t-1}})$ é a projeção de $\boldsymbol{x_t}$ quando é conhecida a série de variáveis observáveis até a data $t-1$.

$\widehat{\boldsymbol{P}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}})$ é a projeção para a variável observável para a data t quando só são conhecidos os valores da série até $t-1$.

Um raciocínio semelhante deve ser feito para chegar-se à equação de atualização do MSE.

$$\begin{aligned}\boldsymbol{E}\left\{\left[\boldsymbol{Y_3} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_2},\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_3} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_2},\boldsymbol{Y_1})\right]'\right\} = \\ \boldsymbol{E}\left\{\left[\boldsymbol{Y_3} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_3} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\,\boldsymbol{Y_1})\right]'\right\} - \\ \boldsymbol{E}\left\{\left[\boldsymbol{Y_3} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_2} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]'\right\}\end{aligned}$$

$$\times \left(\boldsymbol{E}\left\{\left[\boldsymbol{Y_2} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_2} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]'\right\}\right)^{-1} \mathbf{H_{23}}$$

Mas na equação acima H_{23} pode ser substituído por $\mathrm{H}_{32}{}'$, onde $\mathbf{H_{32}}'$ será igual a

$$\mathbf{H_{32}}' = \mathbf{H_{23}} = \left(\boldsymbol{E}\left\{\left[\boldsymbol{Y_3} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_2} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]'\right\}\right)'$$

$$\mathbf{H_{23}} = \boldsymbol{E}\left\{\left[\boldsymbol{Y_2} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_2}|\boldsymbol{Y_1})\right]\left[\boldsymbol{Y_3} - \widehat{\boldsymbol{P}}(\boldsymbol{Y_3}|\boldsymbol{Y_1})\right]'\right\}$$

Logo, o MSE as relações entre as variáveis sugeridas acima será

$$\boldsymbol{E}\left\{\left[\boldsymbol{x_t} - \widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{z_t}, \boldsymbol{Z_{t-1}})\right]\left[\boldsymbol{x_t} - \widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{z_t}, \boldsymbol{Z_{t-1}})\right]'\right\} =$$

$$\boldsymbol{E}\left\{\left[\boldsymbol{x_t} - \widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{Z_{t-1}})\right]\left[\boldsymbol{x_t} - \widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\ \boldsymbol{Z_{t-1}})\right]'\right\} -$$

$$\boldsymbol{E}\left\{\left[\boldsymbol{x_t} - \widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{Z_{t-1}})\right]\left[\boldsymbol{z_t} - \widehat{\boldsymbol{P}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}})\right]'\right\} \times$$

$$\left(\boldsymbol{E}\left\{\left[\boldsymbol{z_t} - \widehat{\boldsymbol{P}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}})\right]\left[\boldsymbol{z_t} - \widehat{\boldsymbol{P}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}})\right]'\right\}\right)^{-1} \times$$

$$\boldsymbol{E}\left\{\left[\boldsymbol{z_t} - \widehat{\boldsymbol{P}}(\boldsymbol{z_t}|\boldsymbol{Z_{t-1}})\right]\left[\boldsymbol{x_t} - \widehat{\boldsymbol{P}}(\boldsymbol{x_t}|\boldsymbol{Z_{t-1}})\right]'\right\}$$

Alterando as variáveis para aquelas adotas neste texto chega-se a

$$\boldsymbol{P_{t^+}} = \boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})'] - \boldsymbol{E}[(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})']$$
$$\times \{\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})']\}^{-1} \times \{\boldsymbol{E}[(\boldsymbol{z_t} - \boldsymbol{z_{t^-}})(\boldsymbol{x_t} - \boldsymbol{x_{t^-}})']\}$$

Através destas duas equações de atualização termina-se a demonstração das equações do algoritmo do FK.

16.4. Maximização da Função de Verossimilhança Quando as Observações se Distribuem Normalmente

Como a f.d.p. conjunta é, conforme descrição feita durante o texto, dada por:

$$\prod_{t=1}^{T} f(y_t;\ \mu, \sigma^2)$$

Então o logaritmo desta função será dado pelo

$$\sum_{t=1}^{T} f(y_t;\ \mu, \sigma^2)$$

Pois, o logaritmo do produto de duas variáveis é igual a soma dos logaritmos das variáveis. Assim deseja-se maximizar o logaritmo da função de verossimilhança dado por:

$$ln[L(y_t;\ \mu, \sigma^2)] = \sum_{t=1}^{T} ln[f(y_t;\ \mu, \sigma^2)]$$

Mas, substituindo a f.d.p. pela equação (96):

$$ln[L(y_t;\ \mu,\sigma^2)] = \sum_{t=1}^{T} ln\left[\frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2\sigma^2}(y_t-\mu)^2\right)\right]$$

Aplicando a propriedade de que o logaritmo do produto é igual a soma dos logaritmos:

$$ln[L(y_t;\ \mu,\sigma^2)] = \sum_{t=1}^{T} ln\left[\frac{1}{\sqrt{2\pi\sigma^2}}\right] + ln\left[exp\left(-\frac{1}{2\sigma^2}(y_t-\mu)^2\right)\right]$$

O somatório acima pode ser dividido em dois somatórios, tal que:

$$ln[L(y_t;\ \mu,\sigma^2)] = \sum_{t=1}^{T} ln\left[\frac{1}{\sqrt{2\pi\sigma^2}}\right] + \sum_{t=1}^{T} ln\left[exp\left(-\frac{1}{2\sigma^2}(y_t-\mu)^2\right)\right]$$

O primeiro somatório é o somatório de uma constante, e consequentemente será igual a n vezes a constante. No segundo somatório há o logaritmo neperiano de uma exponencial, o qual por definição é igual ao termo ao qual a exponencial está elevada.

$$ln[L(y_t;\ \mu,\sigma^2)] = T ln\left[\frac{1}{\sqrt{2\pi\sigma^2}}\right] + \sum_{t=1}^{T} -\frac{1}{2\sigma^2}(y_t-\mu)^2$$

Ao primeiro termo é possível aplicar a definição de que o logaritmo da razão é igual à subtração dos logaritmos. No segundo termo há uma constante que pode ser retirada do somatório.

$$ln[L(y_t;\ \mu,\sigma^2)] = T\left[ln(1) - ln\left(\sqrt{2\pi\sigma^2}\right)\right] - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t-\mu)^2$$

Ou,

$$ln[L(y_t;\ \mu,\sigma^2)] = -T ln\left[(2\pi\sigma^2)^{\frac{1}{2}}\right] - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t-\mu)^2$$

Desenvolvendo,

$$ln[L(y_t;\ \mu,\sigma^2)] = -\frac{T}{2} ln[2\pi\sigma^2] - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t-\mu)^2$$

Consequentemente,

$$ln[L(y_t;\ \mu,\sigma^2)] = -\frac{T}{2} ln[2\pi] - \frac{T}{2} ln[\sigma^2] - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t-\mu)^2$$

A equação acima representa o logaritmo da função de verossimilhança. Portanto, é a relação a qual se deseja maximizar. A seguir apresentam-se as derivadas da função acima em relação aos parâmetros da distribuição.

$$\frac{\partial ln(L)}{\partial \mu} = -\frac{1}{2\sigma^2}\left(\frac{\partial}{\partial \mu}\sum_{t=1}^{T} {y_t}^2 - \frac{\partial}{\partial \mu}\sum_{t=1}^{T} 2y_t\mu + \frac{\partial}{\partial \mu}\sum_{t=1}^{T} \mu^2\right)$$

Desenvolvendo

$$\frac{\partial ln(L)}{\partial \mu} = -\frac{1}{2\sigma^2}\left(-2\sum_{t=1}^{T} y_t + 2\sum_{t=1}^{T} \mu\right)$$

Multiplicando

$$\frac{\partial ln(L)}{\partial \mu} = \frac{1}{\sigma^2}\sum_{t=1}^{T} y_t - \frac{1}{\sigma^2}T\mu$$

Como se deseja maximizar

$$\frac{\partial ln(L)}{\partial \mu} = \frac{1}{\sigma^2}\sum_{t=1}^{T} y_t - \frac{1}{\sigma^2}T\mu \ = 0$$

$$\frac{\partial ln(L)}{\partial \sigma^2} = -\frac{T}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{t=1}^{T}(y_t - \mu)^2 = 0$$

Resolvendo a primeira equação acima para μ:

$$\frac{1}{\sigma^2}\sum_{t=1}^{T} y_t - \frac{1}{\sigma^2}T\mu \ = 0$$

Isolando o termo com μ

$$\frac{1}{\sigma^2}\sum_{t=1}^{T} y_t = \frac{1}{\sigma^2}T\mu$$

Simplificando:

$$\sum_{t=1}^{T} y_t = T\mu$$

Então

$$\hat{\mu} = \frac{\sum_{t=1}^{T} y_t}{T}$$

Resolvendo a equação, referente a derivação da verossimilhança em função de σ^2, para σ^2:

$$-\frac{T}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{t=1}^{T}(y_t - \mu)^2 = 0$$

Logo,

$$\frac{1}{2(\sigma^2)^2}\sum_{t=1}^{T}(y_t - \mu)^2 = \frac{T}{2}\frac{1}{\sigma^2}$$

Simplificando

$$\frac{\sum_{t=1}^{T}(y_t-\mu)^2}{\sigma^2} = T$$

Isolando σ^2:

$$\widehat{\sigma^2} = \frac{\sum_{t=1}^{T}(y_t-\mu)^2}{T}$$

17
Apêndice H – Demonstração do Lema de Inversão de Matrizes (Álgebra)

Esse Lema pode ser facilmente demonstrado. O método utilizado aqui para encontrar-se a validade da equação será o de partir dela e chagar a uma igualdade trivial. O Lema de Inversão de matrizes diz que:

$$(\boldsymbol{A}+\boldsymbol{BC})^{-1}=\boldsymbol{A}^{-1}-\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CA}^{-1}$$

Logo, é possível multiplicar todos os termos da equação pela matriz original, aquela que se deseja inverter, lembrando que quando trabalha-se com matrizes a ordem dos fatores altera o produto, e portanto, a matriz pela qual a equação está sendo multiplicada será inserida sempre antes dos termos da equação.

$$(\boldsymbol{A}+\boldsymbol{BC})(\boldsymbol{A}+\boldsymbol{BC})^{-1}=(\boldsymbol{A}+\boldsymbol{BC})\boldsymbol{A}^{-1}-(\boldsymbol{A}+\boldsymbol{BC})\boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I}++\boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CA}^{-1}$$

O lado esquerdo da equação torna-se igual à identidade, pois sabe-se que

$$\boldsymbol{AA}^{-1}=\boldsymbol{I}$$

Além disso, os dois termos do lado direito são abertos e, desta forma,

$$\boldsymbol{I}=\boldsymbol{I}+\boldsymbol{BCA}^{-1}-\boldsymbol{IB}(\boldsymbol{I}+\boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CA}^{-1}-\boldsymbol{BCA}^{-1}\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CA}^{-1}$$

Passando a matriz identidade do lado direito para o esquerdo, fica-se com uma matriz de zeros daquele lado, e também sabe-se que $\boldsymbol{AI}=\boldsymbol{A}$, e , portanto, a relação se torna

$$\boldsymbol{0}=\boldsymbol{BCA}^{-1}-\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CA}^{-1}-\boldsymbol{BCA}^{-1}\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}\boldsymbol{CA}^{-1}$$

Como a matriz $\boldsymbol{B}$ está em todos os termos podemos sumir com ela multiplicando todos os termos por $\boldsymbol{B}^{-1}$. Posteriormente, basta executar o mesmo procedimento para a matriz $\boldsymbol{CA}^{-1}$, ou seja, multiplicar todos os termos por $(\boldsymbol{CA}^{-1})^{-1}$. Assim, chega-se a

$$\boldsymbol{0}=\boldsymbol{I}-(\boldsymbol{I}+\boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}-\boldsymbol{CA}^{-1}\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}$$

Procedimento semelhante pode ser aplicado para a matriz $(\boldsymbol{I}+\boldsymbol{CA}^{-1}\boldsymbol{B})^{-1}$, multiplicando todos os termos por $(\boldsymbol{I}+\boldsymbol{CA}^{-1}\boldsymbol{B})$

$$\mathbf{0} = (\boldsymbol{I} + \boldsymbol{C}\boldsymbol{A}^{-\mathbf{1}}\boldsymbol{B}) - \boldsymbol{I} - \boldsymbol{C}\boldsymbol{A}^{-\mathbf{1}}\boldsymbol{B}$$

Trabalhando a equação acima chega-se a

$$\boldsymbol{I} = \boldsymbol{I}$$

Ou,

$$\mathbf{0} = \mathbf{0}$$

Portanto, pode-se constatar que a equação enunciada fornece uma igualdade correta, pois como demonstrado acima, a simplificação da mesma remete a igualdade trivial que obviamente é verdadeira.

18
Apêndice I – Gráficos Para Análise do Ajuste do Modelo aos Dados Reais

Neste apêndice são apresentados os gráficos que não foram inseridos no texto. Desta forma, são apresentados os resultados para os contratos F9, F13 e F17. Primeiramente pode-se ver os gráficos das séries reais e filtradas, e posteriormente são apresentados os resíduos do modelo.

18.1. Gráficos de Ajuste dos Contratos Futuros Observados vs Filtrados

O gráfico abaixo apresenta o comportamento das séries observada e filtrada do contrato F9.

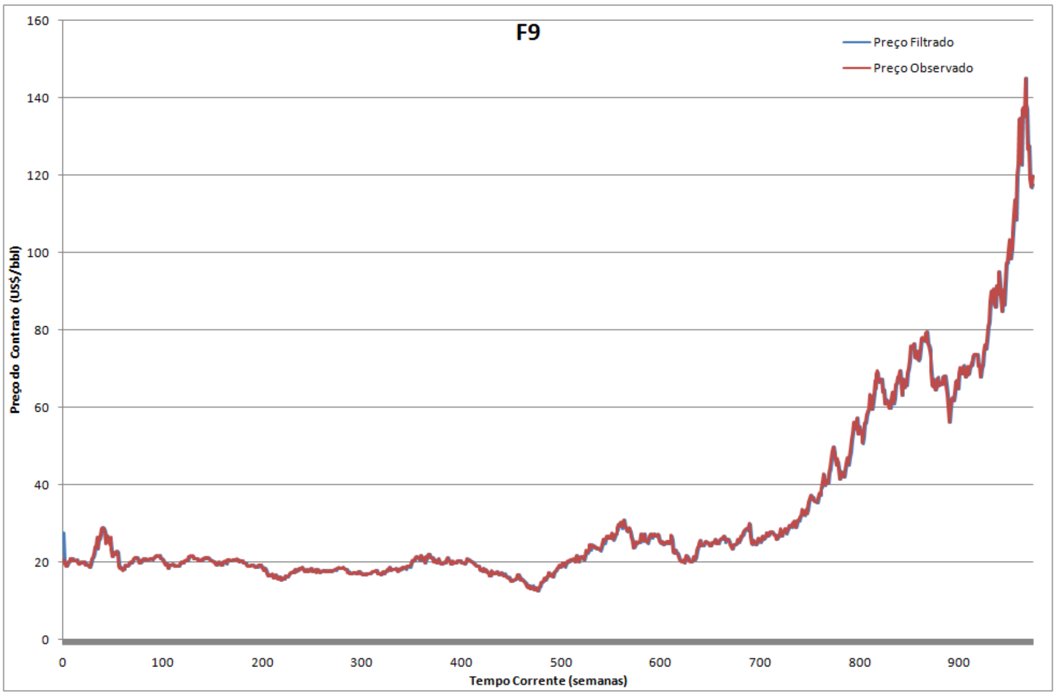

Figura 20 - Preço Observado vs Filtrado Para o Contrato F9

A análise do gráfico acima, da mesma forma do que a dos contratos F1 e F5, permite a conclusão de que a aderência é bastante significativa também para o contrato F9. O único ponto que parece apresentar uma discrepância maior entre o valor filtrado e o observado é exatamente o primeiro. Essa diferença ocorre por conta da ausência de dados para que o filtro faça uma estimativa mais correta. No entanto, pode-se constatar que rapidamente a série de dados filtrados converge para a observada, demonstrando que o ajuste é adequado quando se detém uma observação pelo menos.

O próximo gráfico apresenta as séries filtrada e observada para o contrato F13.

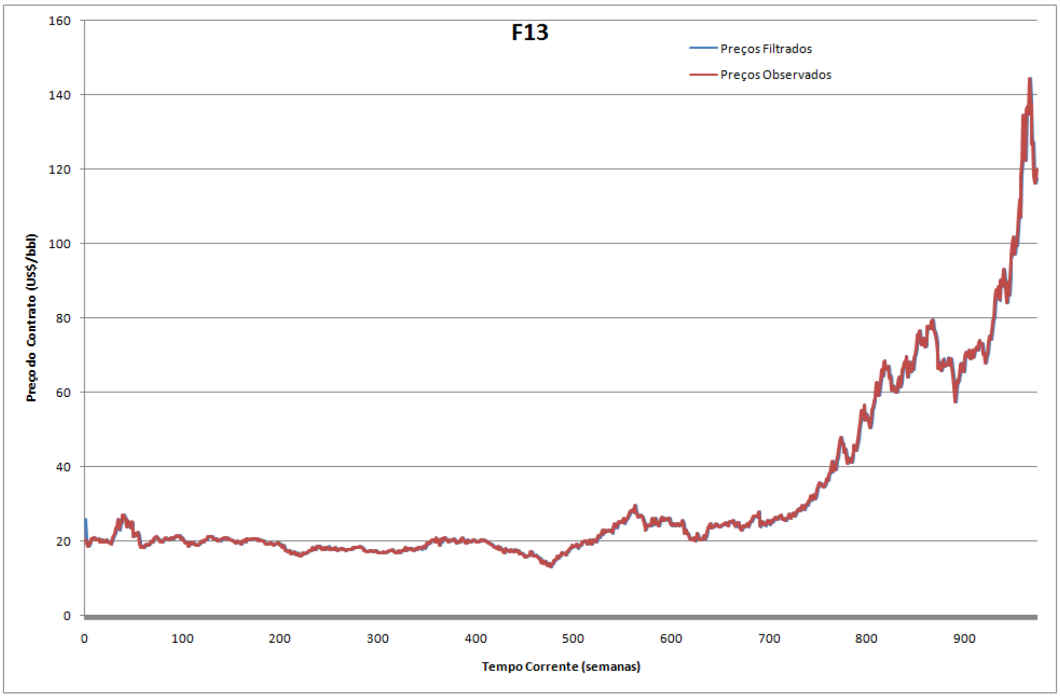

Figura 21 - Preço Observado vs Filtrado Para o Contrato F13

O contrato F13 apresentou excelente ajuste com o modelo estimado pelo Filtro de Kalman, quando a análise é feita através do gráfico acima. O único ponto em que houve discrepância notável pelo "*plot*" acima foi o primeiro, sendo a justificativa idêntica a dada anteriormente para o contrato F9.

A seguir apresenta-se o último gráfico para análise do ajuste visual entre o contrato futuro observado no mercado e o valor filtrado.

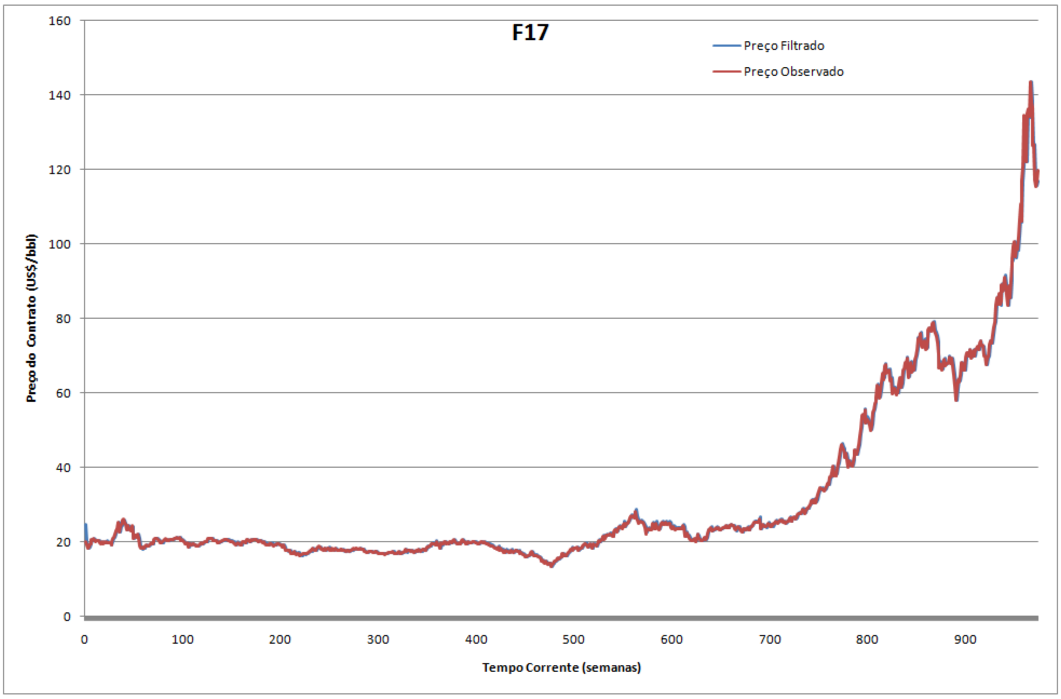

Figura 22 - Preço Observado vs Filtrado Para o Contrato F17

Por último, a análise das curvas do preço observado e do preço filtrado acima permite a constatação da aparente sobreposição das séries.

18.2. Gráficos de Resíduos

Abaixo os resíduos dos contratos F9.

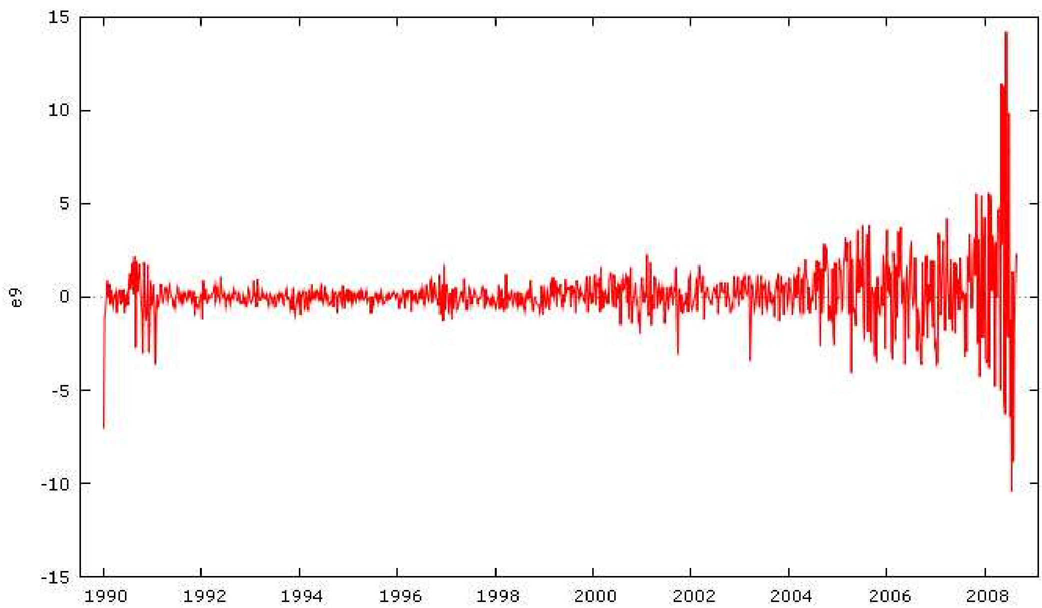

Figura 23 - Erros de Estimativa Para o Contrato F9

O gráfico acima também apresenta aumento da volatilidade dos resíduos a partir de 2004, período em que os preços apresentaram forte tendência de crescimento.

O contrato F13 apresentou resíduos conforme gráfico abaixo.

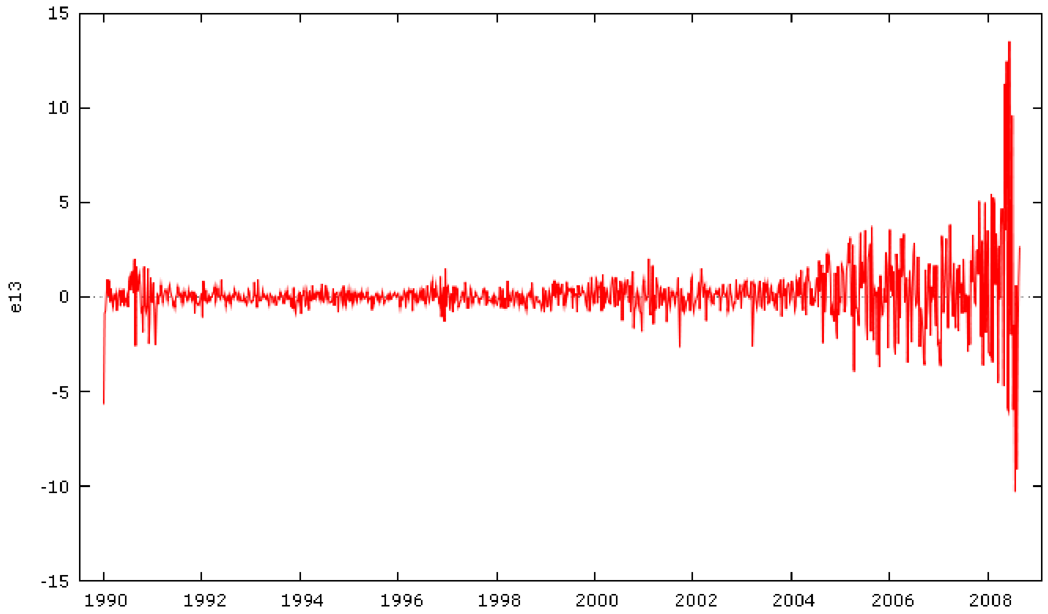

Figura 24 - Erros de Estimativa Para o Contrato F13

Os erros de previsão para o contrato F13 foram bem reduzidos para até 2004, e conforme o comportamento dos outros resíduos aumentaram a partir daquela data.

A seguir é apresentado o gráfico dos resíduos referentes ao contrato F17.

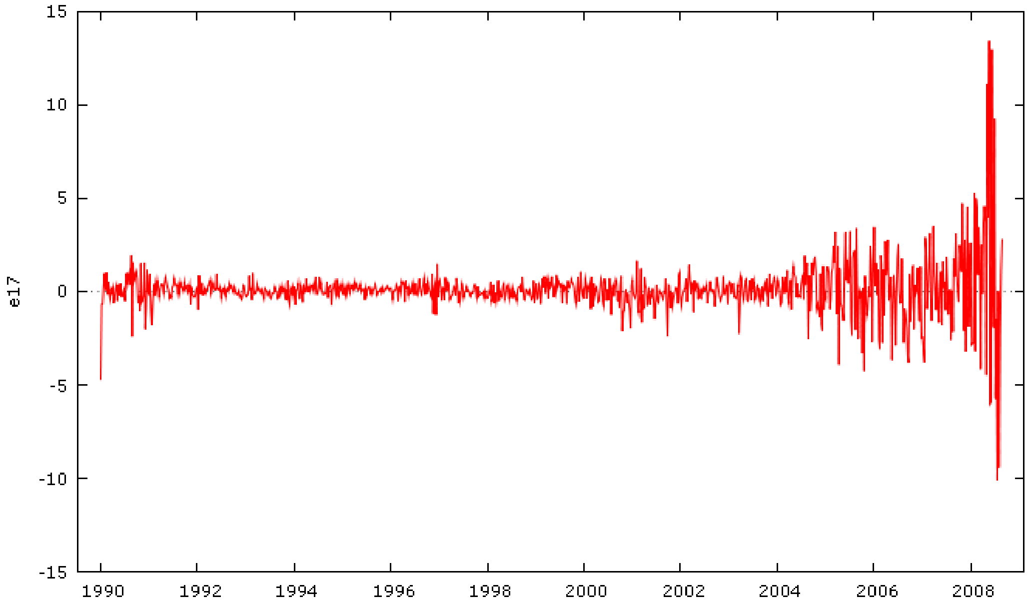

Figura 25 - Erros de Estimativa Para o Contrato F17

18.3. Análise qualitativa da Normalidade dos Resíduos

Os resíduos entre a série observada e a filtrada do contrato F9 encontram-se no gráfico abaixo.

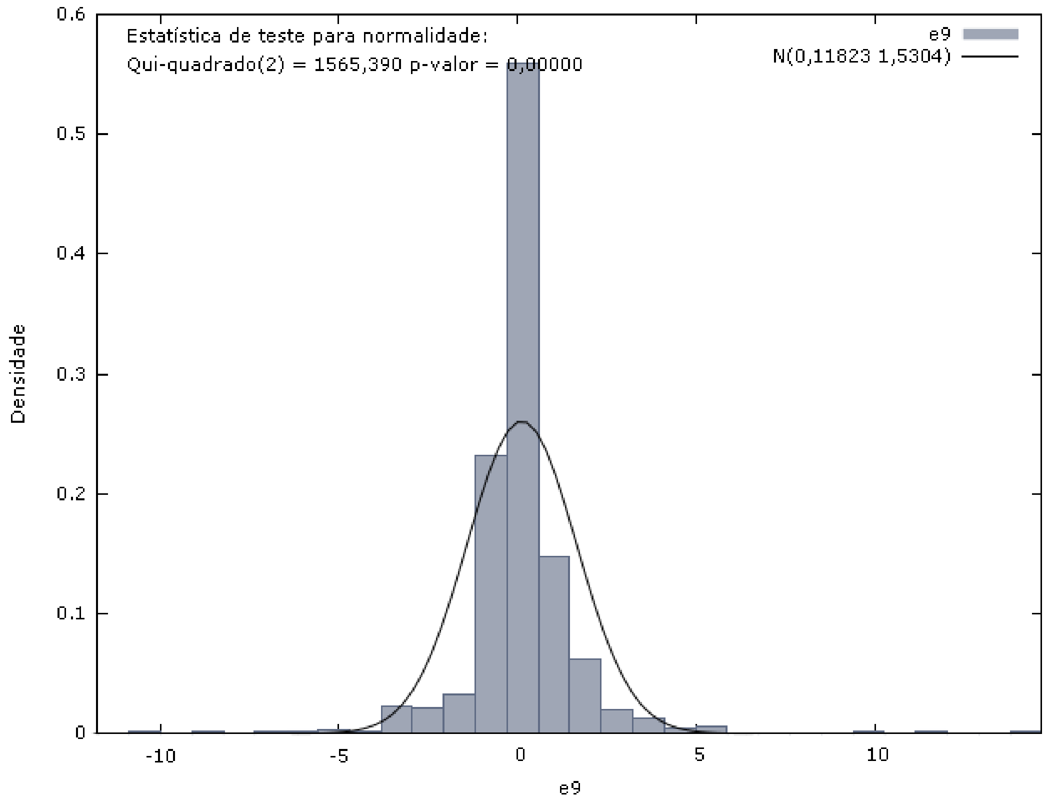

Figura 26 - Distribuição dos Resíduos do Contrato F9

Para o contrato F9 os resíduos se distribuíram semelhantemente aos dos contratos F1 e F5.

A seguir apresentam-se os resíduos do contrato F13.

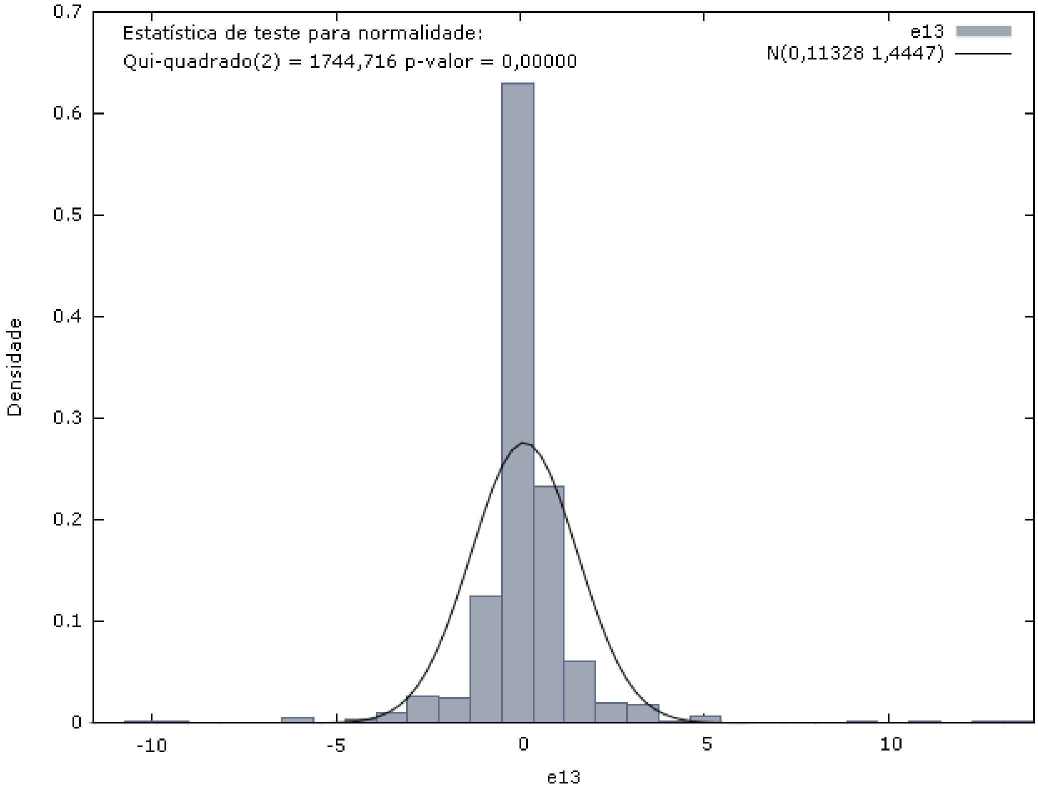

Figura 27 - Distribuição dos Resíduos do Contrato F13

Finalmente, para o contrato F17 as diferenças entre a série observada e filtrada se distribuíram conforme o gráfico de freqüências a baixo.

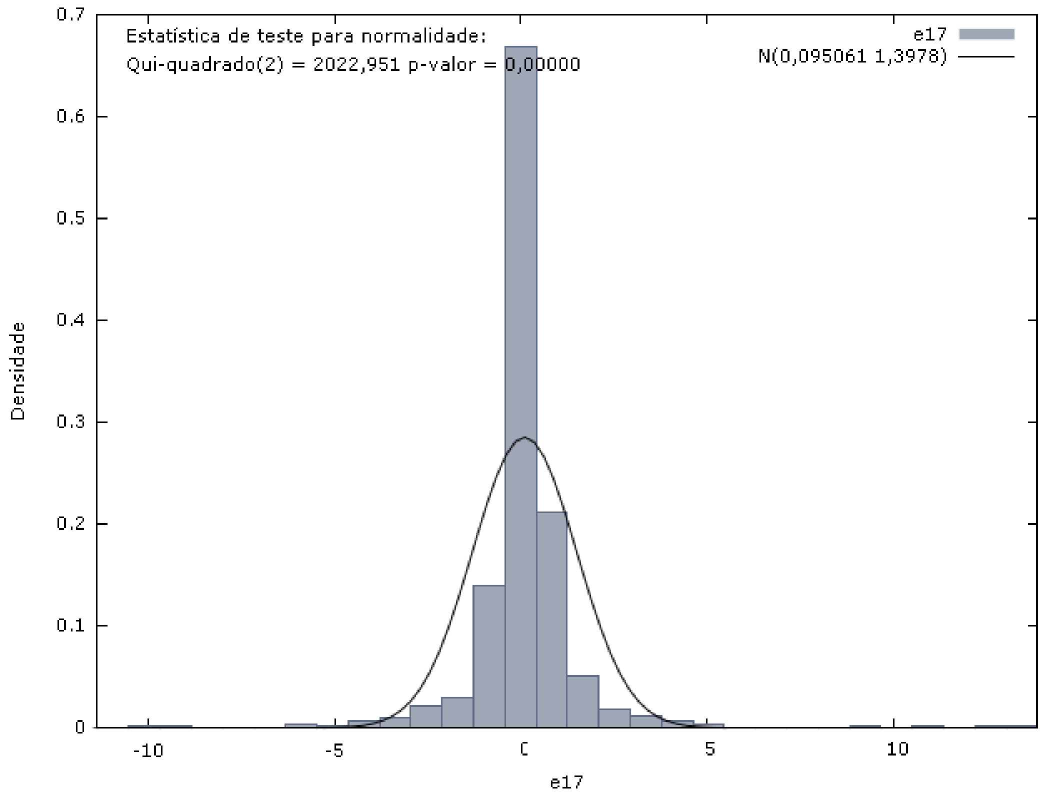

Figura 28 - Distribuição dos Resíduos do Contrato F17