Referências Bibliográficas

[1] ALLAHVIRANLOO, T.; HAJARI, T. **Numerical methods for approximation of fuzzy data**. Applied Mathematics and Computations, 169(1):16–33, 2005.

[2] ANDERSON, T. W.; TAYLOR, J. B. **Strong consistency of least squares estimates in normal linear regressions**. The Annals of Statistics, 4(4):788–790, 1976.

[3] BORDIGNON, A.; CASTRO, R.; LEWINER, T.; LOPES, H.; MURAD, A.; SANTOS, R. ; TAVARES, G. **Métodos baseados em núcleos e máquinas de suporte vetorial em aplicações de geofísica de petróleo**. In: 10TH INTERNATIONAL CONGRESS OF THE BRAZILIAN GEOPHYSICAL SOCIETY, 2007.

[4] BORDIGNON, A.; CASTRO, R.; LEWINER, T.; LOPES, H. ; TAVARES, G. **Point set compression through BSP quantization**. In: SIBGRAPI, p. 229–236. IEEE, 2006.

[5] BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A. ; STONE, C. J. **Classification and Regression Trees**. Chapman & Hall, 1984.

[6] CASELLA, G.; BERGER, R. L. **Statistical Inference**. Duxbury, 2002.

[7] CASTRO, R.; BORDIGNON, A.; LOPES, H.; LEWINER, T. ; TAVARES, G. **Multi-attribute seismic cell reservoir/non-reservoir classification with kernel based methods**. In: 27TH GOCAD MEETING, 2007.

[8] CHANG, C.-C.; LIN, C.-J. **LIBSVM: a library for support vector machines**, 2001. Programa disponível no endereço `http://www.csie.ntu.edu.tw/ cjlin/libsvm`.

[9] CHAUDHURI, P.; HUANG, M.-C.; LOH, W.-Y. ; YAO, R. **Piecewise-polynomial regression trees**. Statistica Sinica, 4(1):143–167, 1994.

[10] CRISTIANINI, N.; SHAWE-TAYLOR, J. **An Introduction to Support Vector Machines and Other Kernel-based Learning Methods**. Cambridge University Press, 2000.

[11] DA ROSA, J. C.; VEIGA, A. ; MEDEIROS, M. C. **Tree-structured smooth transition regression models**. Computational Statistics & Data Analysis, 52:2469–2488, 2008.

[12] DE JONG, P. **Forecasting general insurance liabilities**. Department of Actuarial Studies Research Series - Division of Economic and Financial Studies, Macquarie University, 2004.

[13] DORAY, L. G. **Umvue of the IBNR reserve in a lognormal linear regression model**. Insurance: Mathematics and Economics, 18(1):43–57, 1996.

[14] EICKER, F. **Asymptotic normality and consistency of the least squares estimators for families of linear regressions**. The Annals of Mathematical Statistics, 34(2):447–456, 1963.

[15] ENGLAND, P. D.; VERRALL, R. J. **Stochastic claims reserving in general insurance**. British Actuarial Journal, 8(3):443–518(76), 2002.

[16] FAN, R.-E.; CHEN, P.-H. ; LIN, C.-J. **Working set selection using the second order information for training SVM**. Journal of Machine Learning Research, 6:1889–1918, 2005.

[17] FRIEDMAN, J. H. **Multivariate adaptive regression splines**. The Annals of Statistics, 19(1):1–67, 1991.

[18] GESMANN, M. **Mack-, Bootstrap and Munich-chain-ladder methods for insurance claims reserving**, 2009. R Package at `http://code.google.com/p/chainladder/`.

[19] HASTIE, T.; TIBSHIRANI, R. **Generalized Additive Models**. Monographs on Statisticas and Applied Probability 43. Chapman & Hall, 1990.

[20] HASTIE, T.; TIBSHIRANI, R. ; FRIEDMAN, J. **The Elements of Statistical Learning**. Springer Series in Statistics. Springer, 2001.

[21] HERBST, T. **An application of randomly truncated data models in reserving IBNR claims**. Insurance: Mathematics and Economics, 25(2):123–131, 1999.

[22] HOEL, P. G.; PORT, S. C. ; STONE, C. J. **Introduction to Statistical Theory**. The Houghton-Mifflin series in statistics. Houghton Mifflin, Boston, 1971.

[23] HU, C.; CARDENAS, A.; HOOGENDOORN, S. ; JR., P. S. **An interval polynomial interpolation problem and its lagrange solution**. Reliable Computing, 4(1):27–38, 1998.

[24] KREINOVICH, V. **Soft Methods for Integrated Uncertainty Modelling**, volume 37/2006 de **Advances in Soft Computing**, capítulo Statistical Data Processing under Interval Uncertainty: Algorithms and Computational Complexity, p. 11–26. Springer Berlin, Heidelberg, 2006.

[25] KUBRUSLY, J.; DIAS, M.; LOPES, H. ; LAGE, M. **A hybrid chain ladder and gaussian process regression method for IBNR estimation**. Contributed Paper at The Fourth Brazilian Conference On Statistical Modelling in Insurance and Finance, 2009.

[26] KUBRUSLY, J.; LOPES, H. ; VEIGA, A. **A probabilistic IBNR estimator using policy data**. Abstracts of the 10th International conference on Insurance: Mathematics and Economics, 1:15–15, 2006.

[27] KUBRUSLY, J.; LOPES, H. ; VEIGA, A. **Um método probabilístico para cálculo de reservas do tipo IBNR**. Revista Brasileira de Risco e Seguro, 4(7):17–46, 2008.

[28] KUBRUSLY, J.; VEIGA, A. ; LOPES, H. **Yet another IBNR estimator**. Proceedings of the Second Brazilian Conference on Statistical Modelling in Insurance and Finance, 2005.

[29] KUZMAN, H. **A support vector machine for AVO interpretation**. 2003 SEG Abstracts, p. 181–184, 2003.

[30] LAGE, M.; BORDIGNON, A.; PETRONETTO, F.; VEIGA, A.; TAVARES, G.; LEWINER, T. ; LOPES, H. **Approximations by smooth transitions in binary space partitions**. In: SIBGRAPI, p. 230–236. IEEE, 2008.

[31] LI, J. **Multiatributes pattern recognition for reservoir prediction**. Geophys. Innovation: CSEG National Convention, p. 205–208, 2005.

[32] LI, J.; CASTAGNA, J. **Support vector machine (svm) pattern recognition to AVO classification**. Geophysical Research Letters, 31(L02609), 2004.

[33] LI, Q.; GRIFFITHS, J. ; WARD, J. **Constructive implicit fitting**. Computer Aided Geometric Design, 23:17–44, 2006.

[34] LI, Q.; TIAN, J. **Blending implicit shapes using fuzzy set operations**. WSEAS Transactions on Information Science and Applications, 5(7):1230–1240, 2008.

[35] MACK, T. **Measuring the variability of chain ladder reserve estimates**. Proceedings of the Casualty Actuarial Society Spring Forum, p. 101–182, 1994.

[36] MACK, T. **Which stochastic model is underlying the chain ladder method?** Insurance: Mathematics and Economics, 15(2-3):133–138, 1994.

[37] MARKOV, S.; POPOVA, E.; SCHNEIDER, U. ; SCHULZE, J. **On linear interpolation under interval data**. Mathematics and Computers in Simulation, 42(1):35–45, 1996.

[38] MOORE, R. E. **Interval Arithmetic and Automatic Error Analysis in Digital Computing**. Tese de Doutorado, Stanford University, October 1962.

[39] MOORE, R. E. **Interval Analysis**. Prentice-Hall, 1966.

[40] MOORE, R. E.; BIERBAUM, F. **Methods and Applications of Interval Analysis (SIAM Studies in Applied and Numerical Mathematics) (Siam Studies in Applied Mathematics, 2.** Soc for Industrial & Applied Math, 1979.

[41] MULANSKY, B.; SCHMIDT, J. W. **Convex interval interpolation using a three-term staircase algorithm**. Numerische Mathematik, 82(2):313–337, 1999.

[42] SHAWE-TAYLOR, J.; CRISTIANINI, N. **Kernel Methods for Pattern Analysis**. Cambridge University Press, 2004.

[43] SMIRNOFF, A.; BOISVERT, E. ; PARADIS, S. J. **Support vector machine for 3d modelling from sparse geological information of various origins**. Computers & Geosciences, 34(2):127–143, 2008.

[44] SMOLA, A. J.; SCHÖLKOPF, B. ; OLKOPF, B. S. **A tutorial on support vector regression**. Technical report, Statistics and Computing, 2003.

[45] STOLFI, J.; DE FIGUEIREDO, L. H. **Self-Validated Numerical Methods and Applications**. Course in the 21st Brazilian Mathematics Colloquium, IMPA, 1997.

[46] TARBELL, T. F. **Incurred but not reported claim reserves**. Proceedings of the Casualty Actuarial Society, Part II Volume XX, 1934.

[47] VAPNIK, V. N. **The Nature of Statistical Learning Theory**. Statistics for Engineering and Information Science. Springer, 2ed edition, 2000.

[48] WASSERMAN, L. **All of Nonparametric Statistics**. Springer Texts in Statistics. Springer, 2006.

[49] ZHAO, B.; WEI ZHOU, H. **Nonlinear classification of AVO attributes using SVM**. AGU Joint Assembly - New Orleans, 2005.

A
Estimadores por Mínimos Quadrados para Regressões Lineares

Sejam $Y = \{y_1, y_2, \dots, y_N\} \subset \mathbb{R}$ e $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \subset \mathbb{R}^n$ dados da amostra. Suponha que os dados se relacionam da seguinte forma:

$$y_i = \mathbf{f}(\mathbf{x}_i)\beta + \varepsilon_i \ , \quad \forall \ \ 1 \le i \le N \tag{A-1}$$

onde $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ é uma função conhecida, $\beta \in \mathbb{R}^m$ é o vetor de parâmetros da regressão e ε_i é uma variável aleatória normal de média nula e variância σ^2, que representa o erro na estimativa.

Suponha $\mathbf{f}(\mathbf{x}) = (\mathbf{f}_1(\mathbf{x}), \dots, \mathbf{f}_m(\mathbf{x}))$. A equação A-1 pode ser escrita na seguinte forma matricial.

$$\underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}}_{y} = \underbrace{\begin{pmatrix} \mathbf{f}_1(\mathbf{x}_1) & \mathbf{f}_2(\mathbf{x}_1) & \dots & \mathbf{f}_m(\mathbf{x}_1) \\ \mathbf{f}_1(\mathbf{x}_2) & \mathbf{f}_2(\mathbf{x}_2) & \dots & \mathbf{f}_m(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \mathbf{f}_1(\mathbf{x}_N) & \mathbf{f}_2(\mathbf{x}_N) & \dots & \mathbf{f}_m(\mathbf{x}_N) \end{pmatrix}}_{\mathbf{M}} \underbrace{\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}}_{\beta} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{pmatrix}}_{\varepsilon} \tag{A-2}$$

O estimador por mínimos quadrados para β é aquele que minimiza a soma dos quadrados dos erros. Ou seja,

$$\widehat{\beta} = \arg \min_{\beta \in \mathbb{R}^m} \sum_{i=1}^{N} (y_i - \mathbf{f}(\mathbf{x}_i)\beta)^2 = \arg \min_{\beta \in \mathbb{R}^m} \sum_{i=1}^{N} \left(y_i - \sum_{j=1}^{m} \mathbf{f}_j(\mathbf{x}_i)\beta_j \right)^2$$

Para encontrar $\widehat{\beta}$ a equação acima será derivada e igualada a zero. Como resultado tem-se:

$$\frac{\partial}{\partial \beta_k} = 0 \quad \Rightarrow \quad \sum_{i=1}^{N} \sum_{j=1}^{m} \mathbf{f}_k(\mathbf{x}_i)\mathbf{f}_j(\mathbf{x}_i)\beta_j = \sum_{i=1}^{N} \mathbf{f}_k(\mathbf{x}_i) y_i$$

Considerando as derivadas em todas as variáveis, $\beta_1 \dots \beta_m$, o problema passa a ser um sistema $m \times m$ representado pela seguinte forma matricial:

$$\mathbf{M}^t\mathbf{M}\beta = \mathbf{M}^t y \tag{A-3}$$

onde a matriz $\mathbf{M}$ e os vetores β e y estão definidos na equação A-2.

A.1 Condição para Existência

Para que exista $\widehat{\beta}$ que seja solução do sistema A-3 é suficiente que a matriz $\mathbf{M}^t\mathbf{M}$ seja inversível. Nesse caso:

$$\widehat{\beta} = \left(\mathbf{M}^t\mathbf{M}\right)^{-1} \mathbf{M}^t y \tag{A-4}$$

A.2 Estimador não Tendencioso

Definição A.2.1 *Um estimador $\widehat{\beta}$ é chamado um estimador não tendencioso para o parâmetro β se $\mathbf{E}[\widehat{\beta}] = \beta$.*

A definição acima foi tirada do livro do Hoel, Port e Stone (22). Observe que $\mathbf{E}[\widehat{\beta}] = \beta$ é o mesmo que $\mathbf{E}[\widehat{\beta} - \beta] = 0$.

Proposição A.2.2 *Supondo $\mathbf{M}^t\mathbf{M}$ não singular, é possível afirmar que o estimador $\widehat{\beta}$ definido pela equação A-4 é não tendencioso.*

Demonstração:

$$\begin{aligned}
\widehat{\beta} &= \left(\mathbf{M}^t\mathbf{M}\right)^{-1} \mathbf{M}^t y \\
\widehat{\beta} &= \left(\mathbf{M}^t\mathbf{M}\right)^{-1} \mathbf{M}^t \left(\mathbf{M}\beta + \varepsilon\right) \\
\widehat{\beta} &= \left(\mathbf{M}^t\mathbf{M}\right)^{-1} \mathbf{M}^t\mathbf{M}\beta + \left(\mathbf{M}^t\mathbf{M}\right)^{-1} \mathbf{M}^t\varepsilon \\
\widehat{\beta} &= \beta + \left(\mathbf{M}^t\mathbf{M}\right)^{-1} \mathbf{M}^t\varepsilon \\
\widehat{\beta} - \beta &= \left(\mathbf{M}^t\mathbf{M}\right)^{-1} \mathbf{M}^t\varepsilon \\
&\Downarrow \\
\mathbf{E}\left[\widehat{\beta} - \beta\right] &= \mathbf{E}\left[\left(\mathbf{M}^t\mathbf{M}\right)^{-1} \mathbf{M}^t\varepsilon\right] = 0
\end{aligned}$$

□

A.3 Condição para Consistência

Observe que $\widehat{\beta}$ depende do tamanho da amostra, uma vez que $\mathbf{M}$ e y dependem de N. Por isso, nessa seção, ele será chamado de $\widehat{\beta}_N$ e $\mathbf{M}^t\mathbf{M}$ será chamado de $\mathbf{M}^t\mathbf{M}_N$. Assim $\{\widehat{\beta}_N\}_{N=1}^{\infty}$ define uma sequência de estimadores.

A seguir, a definição de estimador consistente, retirada do livro do Casella e Berger (6).

Definição A.3.1 *Uma sequência se estimadores* $\widehat{\beta}_N$ *é uma sequência consistente de estimadores para o parâmetro* β *se* $\widehat{\beta}_N$ *converge em probabilidade para* β*. Isto é, se para todo* $\epsilon > 0$*,* $\lim_{N\to\infty} P\left(|\beta_N - \beta| < \epsilon\right) = 1$.

Já foi mostrado, por Anderson e Taylor em (2) e por Eicker em (14), que a condição necessária e suficiente para $\widehat{\beta}_N$ ser um estimador consistente para β é que $\lambda_{\min}\left(\mathbf{M}^{\mathbf{t}}\mathbf{M}_N\right) \xrightarrow[N\to\infty]{} \infty$, onde $\lambda_{\min}\left(\mathbf{M}^{\mathbf{t}}\mathbf{M}_N\right)$ é o menor autovalor da matriz $\mathbf{M}^{\mathbf{t}}\mathbf{M}_N$.

B
SVR - Support Vector Regression

Suponha o problema de aprender a relação entre certos pontos $\mathbf{x}_i \in \mathbb{R}^d$ e $y_i \in \mathbb{R}$ a partir de uma amostra de tamanho N definida pelo conjunto $\mathcal{S} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_N, y_N)\}$. Uma das maneiras de resolver esse problema é através da ferramente de aprendisagem computaciobnal SVR.

A idéia principal por trás do SVR consiste em mapear $\mathbf{x} \in \mathbb{R}^d$ em um espaço de dimensão maior que d, chamado de $\mathcal{F}$, através de uma função não linear $\phi : \mathbb{R}^d \to \mathcal{F}$. Depois disso, com os dados já mapeados, é feita uma regressão linear nos dados $\{(\phi(\mathbf{x}_1), y_1), (\phi(\mathbf{x}_2), y_2), \ldots, (\phi(\mathbf{x}_N), y_N)\}$. A função de previsão dessa regressão linear pode ser modelada por

$$f(\mathbf{x}) = \langle \mathbf{w}, \phi(\mathbf{x}) \rangle + b \tag{B-1}$$

onde $b \in \mathbb{R}$ e $\mathbf{w} \in \mathcal{F}$. Observe que encontrar o valores de b e $\mathbf{w}$ é um problema de regressão linear em $\mathcal{F}$, mas não em $\mathbb{R}^d$, por causa da função ϕ.

B.1
Problema Primal

Para determinar os valores de b e $\mathbf{w}$ será resolvido um problema de otimização quadrática, que busca minimizar a soma dos dos erros que ultrapassarem um certo valor. Esse problema é definido por:

$$\begin{array}{rll} \text{Minimize}_{\mathbf{w},b,\xi_i,\hat{\xi}_i} & \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}(\xi_i + \hat{\xi}_i) & \\ \text{sujeito a:} & & \\ & (\langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle + b) - y_i \leq \varepsilon + \xi_i & i = 1, \ldots, N \\ & y_i - (\langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle + b) \leq \varepsilon + \hat{\xi}_i & i = 1, \ldots, N \\ & \xi_i, \hat{\xi}_i \geq 0 & i = 1, \ldots, N \end{array} \tag{B-2}$$

onde as constantes C e ε são definidas pelo usuário e C representa a penalidade realizada quando o erro em um ponto qualquer for maior ε, que indica o erro máximo permitido sem penalidade. Já ξ_i e $\hat{\xi}_i$ são variáveis *slack*, que indicam o quanto o erro no ponto i ultrapassou ε.

B.2
Problema Dual

A forma dual do problema de otimização definido na equação (B-2) pode ser definida por:

$$
\begin{array}{rl}
\text{Maximize}_{\alpha_i,\hat{\alpha}_i} & \sum_{i=1}^{N}(\hat{\alpha}_i - \alpha_i)y_i - \varepsilon\sum_{i=1}^{N}(\hat{\alpha}_i + \alpha_i) - \\
 & \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\hat{\alpha}_i - \alpha_i)(\hat{\alpha}_j - \alpha_j)\langle\phi(\mathbf{x}_i),\phi(\mathbf{x}_j)\rangle \\
\text{sujeito a:} & \\
 & \sum_{i=1}^{N}(\hat{\alpha}_i - \alpha_i) = 0 \quad i = 1,\ldots,N \\
 & 0 \le \alpha_i, \hat{\alpha}_i \le C \qquad i = 1,\ldots,N
\end{array}
\tag{B-3}
$$

onde α_i e $\hat{\alpha}_i$ são os multiplicadores de Lagrange. As condições de Karush-Kuhn-Tucker (KKT) para esse problema são:

$$
\begin{array}{ll}
\alpha_i(\langle\mathbf{w},\phi(\mathbf{x}_i)\rangle + b - y_i - \varepsilon - \xi_i) = 0 & i = 1,\ldots,N \\
\hat{\alpha}_i(y_i - \langle\mathbf{w},\phi(\mathbf{x}_i)\rangle - b - \varepsilon - \hat{\xi}_i) = 0 & i = 1,\ldots,N \\
\hat{\alpha}_i\alpha_i = 0, \hat{\xi}_i\xi_i = 0 & i = 1,\ldots,N \\
(\hat{\alpha}_i - C)\hat{\xi}_i = 0, (\alpha_i - C)\xi_i = 0 & i = 1,\ldots,N
\end{array}
$$

B.3
Núcleo

Na forma dual, exposto na equação (B-3), o problema passa a depender apenas de $\langle\phi(\mathbf{x}_i),\phi(\mathbf{x}_j)\rangle$ e não mais de ϕ propriamente dito. De fato, para resolver esse problema, a função ϕ não será conhecida. O que se conhece é $K(\mathbf{x}_i,\mathbf{x}_j) = \langle\phi(\mathbf{x}_i),\phi(\mathbf{x}_j)\rangle$, que será chamada de função núcleo, ou simplesmente núcleo.

O núcleo é mais um parâmetro escolhido pelo usuário. Existem diversas alternativas, como por exemplo, o núcleo gaussiano definido por

$$K(\mathbf{x}_i,\mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{2\gamma^2}\right).$$

Para conhecer mais sobre os núcleos, sua definição e propriedades, uma boa referência é o livro de Shawe-Taylor e Cristianini (42).

B.4
Solução

Voltando ao problema dual definido na equação (B-3), se $(\hat{\alpha}_i - \alpha_i)$ for substituído por δ_i, utilizando as condições de KKT o problema pode ser reescrito para:

$$\begin{array}{rl} \text{Maximize}_{\delta_1,\ldots,\delta_l} & \sum_{i=1}^{l} \delta_i y_i - \varepsilon \sum_{i=1}^{l} |\delta_i| - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \delta_i \delta_j K(\mathbf{x}_i, \mathbf{x}_j) \\ \text{sujeito a:} & \\ & \sum_{i=1}^{l} \delta_i = 0 \quad i = 1, \ldots, l \\ & 0 \leq \delta_i \leq C \quad i = 1, \ldots, l \end{array} \tag{B-4}$$

Seja $\delta^* = (\delta_1^*, \delta_2^*, \ldots, \delta_N^*)$ a solução do problema exposto na equação (B-4), que pode ser obtida através de qualquer método em otimização quadrática, como por exemplo, para o caso do algoritmo LIBSVM (8) foi usado o SMO (*sequential minimal optimization*). A a função de previsão definida pela equação (B-1) será:

$$f(\mathbf{x}) = \sum_{i=1}^{N} \delta_i^* K(\mathbf{x}_i, \mathbf{x}) + b^*,$$

onde $b^* = y_k - \varepsilon + \sum_{i=1}^{N} \delta_i^* K(\mathbf{x}_i, \mathbf{x}_k)$ para qualquer $\mathbf{x}_k$ tal que $0 < \delta_k^* < C$. Para mais detalhes sobre o modelo e o algoritmo do SVR veja (10) e (44).

C
Aritmética Intervalar

A aritmética intervalar é uma extensão da aritmética para segmentos da reta. Um dos primeiros pesquisadores que desenvolveram essa teoria foi Moore, começando por sua tese de doutorado (38) e seguindo por diversas outras publicações, como por exemplo (39) e (40). A partir daí surgiram diversos artigos e livros sobre aritmética intervalar e o assunto virou foco de muita pesquisa, principalmente na área de computação numérica.

C.1
Intervalos

Um intervalo compacto real, ou somente um intervalo, é um subconjunto não-vazio, fechado e limitado de números reais da forma

$$[x] \equiv [\ \underline{x}\ ,\ \overline{x}\ ] := \{x \in \mathbb{R} \mid \underline{x}\ \leq x \leq\ \overline{x}\ \},$$

onde $\underline{x}$, é o limite inferior e $\overline{x}$ é o limite superior do intervalo.

Os limites inferior e superiror também são chamados, respectivamente, de *ínfimo* e *supremo*, e podem ser denotados por $\inf[x]$ e $\sup[x]$. O conjunto de todos os intervalos reais será chamado por $\mathbb{IR}$.

Um intervalo é chamado *pontual* se $\underline{x} = \overline{x}$, nesse caso pode-se escrever simplesmente x. O *interior* de um intervalo $[x]$, denotado por $]x[$ ou $\text{int}[x]$, é definido por:

$$]x[\equiv \text{int}[x] := \{x \in \mathbb{R} \mid \underline{x} < x < \overline{x}\}.$$

As relações de *igualdade* ($=$), *pertinência* ($\in$), *inclusão* ($\subseteq$), *inclusão própria* ($\subset$), *união* ($\cap$) e *interseção* ($\cap$) são extensões da teoria dos conjuntos.

A ordenação no conjunto $\mathbb{IR}$ pode ser feita a partir da ordenação dos números reais da seguinte forma:

$$[x] < [y] \Leftrightarrow x < y \text{ para todo } x \in [x] \text{ e } y \in [y]$$
$$[x] \leq [y] \Leftrightarrow x \leq y \text{ para todo } x \in [x] \text{ e } y \in [y]$$

C.2 Operações Elementares

As operações de soma, subtração, multiplicação e divisão são definidas em $\mathbb{IR}$ da seguinte maneira:

$$\begin{array}{rrcl} \circ: & \mathbb{IR}\times\mathbb{IR} & \to & \mathbb{IR} \\ & ([x],[y]) & \mapsto & [x]\circ[y] = \{x\circ y \mid x\in[x] \text{ e } y\in[y]\} \end{array}$$

onde $\circ \in \{+,-,\cdot,/\}$ representa qualquer uma das quatro operações elementares. Dessa forma, a definição de $[x]/[y]$ é restrita a intervalos com $0\notin[y]$. Fórmulas mais diretas podem ser usadas:

$$\begin{array}{rcl} [x]+[y] & = & [\underline{x}+\underline{y}\ ,\ \overline{x}+\overline{y}] \\ [x]-[y] & = & [\underline{x}-\overline{y}\ ,\ \overline{x}-\underline{y}] \\ [x]\cdot[y] & = & [\ \min\{\underline{xy},\underline{x}\overline{y},\overline{x}\underline{y},\overline{xy}\}\ ,\ \max\{\underline{xy},\underline{x}\overline{y},\overline{x}\underline{y},\overline{xy}\}\ ] \\ [x]/\,[y] & = & [x]\cdot[1/\overline{y}\ ,\ 1/\underline{y}],\ \text{para } 0\notin[y]. \end{array}$$

É interessante observar que essas operações elementares obedecem a relação de *inclusão isotônica*:

$$[x]\subseteq[x']\,,[y]\subseteq[y']\Rightarrow[x]\circ[y]\subseteq[x']\circ[y'],\quad \forall\circ\in\{+,-,\cdot,/\}.$$

Somente algumas propriedades algébricas da aritmética de números reais permanecem válidas na aritmética intervalar. É possível ver, pela definição, que as operações de adição e multiplicação em $\mathbb{IR}$ satisfazem as leis de comutatividade e associatividade. Os elementos neutros da adição e da multiplicação são, respectivamente, os intervalos $[0,0]$ e $[1,1]$.

Um propriedade que geralmente não é satisfeita é a lei da distributividade. Porém, uma forma mais fraca, chamada a lei de *subdistributividade*, pode ser aplicada em $\mathbb{IR}$:

$$\begin{array}{rcl} [x]\cdot([y]\pm[z]) & \subseteq & [x]\cdot[y]\pm[x]\cdot[z], \\ ([y]\pm[z]).[x] & \subseteq & [y].[x]\pm[z].[x]. \end{array}$$

C.3 Funções Intervalares

Um dos grandes problemas da análise intervalar é avaliar a imagem de uma função real em um intervalo. Isto é, dada uma função real φ,

$$\varphi: D\subset\mathbb{R}\to\mathbb{R},$$

e um intervalo $[x] \in D$, deseja-se encontrar $\varphi([x])$ definida por:

$$\varphi([x]) := \{\varphi(x) \mid x \in [x]\}.$$

Se φ for uma função contínua, então $\varphi([x])$ será um intervalo. Assim a função φ pode ser estendida para o conjunto $\mathbb{IR}$ da seguinte forma:

$$\begin{array}{rlll} \widetilde{\varphi}: & \widetilde{D} \subset \mathbb{IR} & \rightarrow & \mathbb{IR} \\ & [x] & \mapsto & \widetilde{\varphi}([x]) = \varphi([x]) \end{array}$$

onde $\widetilde{D}$ representa todos os intervalos em D.

Usando propriedades de monotonicidade das funções reais elementares, pode-se escrever a imagem de algumas funções estendidas para intervalos, como por exemplo:

$$\begin{array}{ll} \varphi([\underline{x}, \overline{x}]) = [\varphi(\underline{x}), \varphi(\overline{x})]\,, & \text{se } \varphi \text{ é não decrescente} \\ \varphi([\underline{x}, \overline{x}]) = [\varphi(\overline{x}), \varphi(\underline{x})]\,, & \text{se } \varphi \text{ é não crescente} \\ \operatorname{abs}([x]) = [\langle [x] \rangle, |[x]|] & \\ [x]^2 = [\langle [x] \rangle^2, |[x]|^2] & \end{array}$$

onde $\langle [x] \rangle$ representa o menor valor absoluto no intervalo $[x]$ e $|[x]|$ o maior valor absoluto.

Mas nem sempre é possível se encontrar uma expressão que defina a imagem de uma função real em um intervalo. Por isso um dos problemas fundamentais da aritmética intervalar é determinar, através de cálculos numéricos, um intervalo fechado que contenha essa imagem e que seja o mais próximo possível dela.

É fácil obter uma estimativa de $f([x])$ simplesmente substituindo x por $[x]$ na expressão que define f, e então avaliar f usando a aritmética intervalar. Assumindo que todas as operações de aritmética intervalar são bem definidas, este tipo de avaliação é chamada de *extensão natural* de f e é denotada por $f_{[\,]}([x])$.

Teorema C.1 ***(Teorema Fundamental da Aritmética Intervalar)***
Sejam $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ uma função contínua, $f_{[\,]}$ a sua extensão natural e $[x] \in D$ um intervalo qualquer. Então, $f([x]) \subseteq f_{[\,]}([x])$.

A igualdade ocorre em situações muitas raras. Mesmo assim, muitas vezes, a extensão natural é o melhor que se pode fazer para estimar $f([x])$. Um estudo mais completo sobre aritmética intervalar pode ser encontrado no livro de Stolfi e Figueiredo (45).