A
Algoritmo de Blahut

A seguir, apresentamos um desenvolvimento sucinto do algoritmo de Blahut para o cálculo da função taxa-distorção. Mais informações podem ser encontradas em [11] e [6].

Desejamos obter a curva da função $R(D)$ para uma fonte discreta sem memória X com distribuição $P(x)$ e medida de distorção $d(x,y)$. Inicialmente, notamos que a minimização na definição da função $R(D)$ dada em (2-10) pode ser reescrita como uma dupla minimização

$$R(D) = \min_{q(y)} \min_{\substack{Q(y|x): \\ \sum P(x)Q(y|x)d(x,y) \leq D}} I(X;Y) \tag{A-1}$$

$$= \min_{q(y)} \min_{\substack{Q(y|x): \\ \sum P(x)Q(y|x)d(x,y) \leq D}} \sum_x \sum_y P(x)Q(y|x) \log \frac{Q(y|x)}{q(y)} \tag{A-2}$$

onde $q(y)$ representa a distribuição de probabilidade de Y e $Q(y|x)$ representa a distribuição condicional de Y dado X.

A essência do algoritmo de Blahut reside em um processo de minimização alternada: fixada uma distribuição $q(y)$ inicial, calcula-se a distribuição $Q(y|x)$ que minimiza a informação mútua, isto é, efetua-se a segunda minimização na definição acima. Como esta minimização possui restrição, utiliza-se um multiplicador de Lagrange λ para incluir a restrição na distorção. Em seguida, para essa distribuição condicional $Q(y|x)$ encontrada, obtém-se agora uma nova distribuição $q(y)$ realizando-se a primeira minimização. O processo então se repete, sendo calculadas alternadamente as funções $Q(y|x)$ e $q(y)$. Uma vez que a informação mútua é sempre reduzida a cada etapa, em alguma iteração o algoritmo deve atingir um ponto de mínimo (mostra-se que o mínimo obtido é sempre um mínimo global [11]). O algoritmo então chega ao fim, retornando os valores de taxa R e de distorção D, os quais correspondem a um ponto no plano $R \times D$. Outros pontos podem ser obtidos repetindo-se o algoritmo para diversos valores de λ, o que permite traçar a curva completa da função $R(D)$ para o problema em questão.

O algoritmo de Blahut está sumarizado no Algoritmo A.1 a seguir.

Algoritmo A.1 (Cálculo da Função Taxa-Distorção)

1) Especifique o valor de λ.
2) Inicialize $q(y) = \frac{1}{|\mathcal{Y}|}$.
3) Calcule
   a) $Q(y|x) = \frac{q(y)e^{-\lambda d(x,y)}}{\sum_y q(y)e^{-\lambda d(x,y)}}$
   b) $q(y) = \sum_x P(x)Q(y|x)$.
4) Repita o passo 3 até atingir a precisão desejada[1] ou um número máximo de iterações.
5) Calcule
   a) $D = \sum_x \sum_y P(x)Q(y|x)d(x,y)$
   b) $R = \sum_x \sum_y P(x)Q(y|x) \log \frac{Q(y|x)}{q(y)}$.
6) Fim. ∎

[1]Em [6], Blahut desenvolve limitantes inferior e superior que auxiliam na obtenção de qualquer precisão especificada.

B
Estimativas de Valores Esperados

Nos capítulos 2, 3 e 4 são apresentados quantizadores cujos projeto e avaliação de desempenho dependem do cálculo de valores esperados. Neste apêndice, descreve-se como pode ser obtida uma estimativa do valor esperado de uma função de um vetor aleatório a partir de uma seqüência de amostras independentes.

Sejam $\boldsymbol{X}$ um vetor aleatório e $\{\boldsymbol{X}_j\}_{j=1}^{n}$ uma seqüência de vetores aleatórios i.i.d. com a mesma distribuição de $\boldsymbol{X}$. Seja a variável aleatória $Y = g(\boldsymbol{X})$, com $E[Y] = m_Y$ e $E[(Y - m_Y)^2] = \sigma_Y^2$, e assuma que $m_Y < \infty$ e $\sigma_Y^2 < \infty$. Se definirmos a variável aleatória

$$\bar{Y} = \frac{1}{n}\sum_{j=1}^{n} g(\boldsymbol{X}_j) \tag{B-1}$$

então é possível mostrar que

$$E[\bar{Y}] = E[Y] = m_Y \qquad \text{e} \qquad \sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{n} \tag{B-2}$$

isto é, para n suficientemente grande, a variável aleatória $\bar{Y}$ se torna constante e igual a $E[Y]$; nesse caso, $E[Y] = E[\,g(\boldsymbol{X})\,]$ pode ser determinado com uma única realização de $\bar{Y}$.[1] Assumindo que se dispõe de uma seqüência de n amostras independentes do vetor aleatório $\boldsymbol{X}$,

$$\mathcal{T} = \{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(j)}, \ldots, \boldsymbol{x}^{(n)}\} \tag{B-3}$$

construímos a variável

$$\bar{y} = \frac{1}{n}\sum_{j=1}^{n} g(\boldsymbol{x}^{(j)}). \tag{B-4}$$

Quando n é suficientemente grande, o desenvolvimento acima nos permite

[1] A lei dos grandes números assegura que $\bar{Y}$ converge para m_Y: em probabilidade, no sentido da média quadrática e com probabilidade 1 [29].

"substituir" (aproximar) o valor de $E[\, g(\boldsymbol{X})\,]$ por $\bar{y}$, ou seja,

$$E[\, g(\boldsymbol{X})\,] \approx \frac{1}{n} \sum_{j=1}^{n} g(\boldsymbol{x}^{(j)}). \tag{B-5}$$

Além disso,

$$\begin{aligned} E[\, g(\boldsymbol{X})\,|\boldsymbol{X} \in A] &= \frac{E[\, g(\boldsymbol{X})\, \mathbb{1}(\boldsymbol{X} \in A)]}{P(\boldsymbol{X} \in A)} = \frac{E[\, g(\boldsymbol{X})\, \mathbb{1}(\boldsymbol{X} \in A)]}{E[\, \mathbb{1}(\boldsymbol{X} \in A)]} \\ &\approx \frac{\sum_{j=1}^{n} g(\boldsymbol{x}^{(j)})\, \mathbb{1}(\boldsymbol{x}^{(j)} \in A)}{\sum_{j=1}^{n} \mathbb{1}(\boldsymbol{x}^{(j)} \in A)}. \end{aligned} \tag{B-6}$$

C
Cálculo Diferencial Matricial

Alguns conceitos do cálculo diferencial matricial são revisados a seguir. Sejam os vetores

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \tag{C-1}$$

e o escalar y. A *derivada de um escalar com relação a um vetor* e a *derivada de um vetor com relação a um vetor* são definidas, respectivamente, como

$$\frac{\partial y}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial y}{\partial x_1} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{pmatrix} \qquad \frac{\partial \boldsymbol{y}^T}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}. \tag{C-2}$$

Sejam $\boldsymbol{A}$ uma matriz $n \times m$, $\boldsymbol{a}$ um vetor $n \times 1$ e $\boldsymbol{C}$ uma matriz $n \times n$, onde $\boldsymbol{A}$, $\boldsymbol{a}$ e $\boldsymbol{C}$ independem de $\boldsymbol{x}$. As seguintes relações são válidas:

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T \boldsymbol{A}) = \boldsymbol{A} \tag{C-3}$$

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T \boldsymbol{a}) = \frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{a}^T \boldsymbol{x}) = \boldsymbol{a} \tag{C-4}$$

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T) = \boldsymbol{I} \tag{C-5}$$

onde $\boldsymbol{I}$ denota a matriz identidade. Além disso, temos

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x}) = (\boldsymbol{C} + \boldsymbol{C}^T)\boldsymbol{x} \tag{C-6}$$

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^T \boldsymbol{x}) = 2\boldsymbol{x}. \tag{C-7}$$

Bibliografia

[1] BERGER, T.. **Rate Distortion Theory: A Mathematical Basis for Data Compression**. Prentice-Hall, New Jersey, 1971.

[2] BERGER, T.; JELINEK, F. ; WOLF, J. K.. **Permutation codes for sources**. IEEE Transactions on Information Theory, 18:160–169, 1972.

[3] BERGER, T.. **Optimum quantizers and permutation codes**. IEEE Transactions on Information Theory, 18:759–765, 1972.

[4] BERGER, T.. **Minimum entropy quantizers and permutation codes**. IEEE Transactions on Information Theory, 28:149–157, 1982.

[5] BIGLIERI, E. M.; ELIA, M.. **Optimum permutation modulation codes and their asymptotic performance**. IEEE Transactions on Information Theory, 22:751–753, 1976.

[6] BLAHUT, R. E.. **Principles and Practice of Information Theory**. Addison-Wesley, New York, 1987.

[7] BREGA, L. S.. **Codificação de fontes utilizando códigos de permutação**. Dissertação de mestrado, Pontificia Universidade Católica do Rio de Janeiro, 2003.

[8] BRUNO, S. V. B.. **Códigos de permutação vetorial**. Trabalho de fim de curso, Pontificia Universidade Católica do Rio de Janeiro, 2004.

[9] CHOU, P. A.; LOOKABAUGH, T. ; GRAY, R. M.. **Entropy constrained vector quantization**. IEEE Transactions on Information Theory, 37:31–42, 1989.

[10] CONWAY, J. H.; SLOANE, N. J. A.. **Sphere Packings, Lattices and Groups**. Springer-Verlag, New York, 1988.

[11] COVER, T. M.; THOMAS, J. A.. **Elements of Information Theory**. Wiley, New York, 1991.

[12] DUNN, J. G.. **The performance of a class of n-dimensional quantizers for a Gaussian source**. In: PROC. SYMP. SIGNAL TRANSMISSION AND PROCESSING, p. 76–81, Columbia Univ., New York, NY, May 1965.

[13] FINAMORE, W.; PEARLMAN, W.. **Optimal encoding of discrete-time continuous-amplitude memoryless sources with finite output alphabets**. IEEE Transactions on Information Theory, 26:144–155, 1980.

[14] FINAMORE, W. A.; BRUNO, S. V. B. ; BREGA, L. S.. **Código de permutação vetorial aplicado à fonte uniforme**. In: SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES, Rio de Janeiro, RJ, outubro 2003.

[15] FINAMORE, W. A.; BRUNO, S. V. B. ; SILVA, D.. **Vector permutation encoding for the uniform sources**. In: PROC. IEEE DATA COMPRESSION CONFERENCE, p. 539, Snowbird, UT, March 2004.

[16] GOLDBERG, A. V.; KENNEDY, R.. **An efficient cost scaling algorithm for the assignment problem**. Mathematical Programming, 71:153–178, 1995.

[17] GOYAL, V. K.; SAVARI, S. A. ; WANG, W.. **On optimal permutation codes**. IEEE Transactions on Information Theory, 47:2961–2971, 2001.

[18] GRAY, R. M.; NEUHOFF, D. L.. **Quantization**. IEEE Transactions on Information Theory, 44:2325–2383, 1998.

[19] HAYKIN, S.. **Adaptive Filter Theory**. Prentice Hall, Upper Saddle River, NJ, 4th edition, 2001.

[20] INGEMARSSON, I.. **Optimized permutation modulation**. IEEE Transactions on Information Theory, 36:1098–1100, 1990.

[21] KARLOFF, J. K.. **Permutation codes for the Gaussian channel**. IEEE Transactions on Information Theory, 35:726–732, 1989.

[22] KIEFFER, J. C.. **A survey of the theory of source coding with a fidelity criterion**. IEEE Transactions on Information Theory, 39:1473–1490, 1993.

[23] KING, B. M.; NEIFELD, M. A.. **Low-complexity maximum-likelihood decoding of shortened enumerative permutation codes for holographic storage**. IEEE Journal on Selected Areas in Communications, 19:783–790, 2001.

[24] KSCHISCHANG, F. R.; PASUPATHY, S.. **Optimal nonuniform signaling for gaussian channels**. IEEE Transactions on Information Theory, 39:913–929, 1993.

[25] LINDE, Y.; BUZO, A. ; GRAY, R. M.. **An algorithm for vector quantizer design**. IEEE Transactions on Communications, 28:269–275, 1980.

[26] LLOYD, S. P.. **Least squares quantization in PCM**. Bell Labs Technical Note, 1957. Also in IEEE Transactions on Information Theory, 28:129–137, 1982.

[27] LU, L.; COHEN, G. ; GODLEWSKI, P.. **Composite permutation coding of speech waveforms**. IEEE International Conference on Acoustics, Speech, and Signal Processing, 11:2359–2362, 1986.

[28] NORDIO, A.; VITERBO, E.. **Permutation modulation for fading channels**. In: PROC. INT. CONF. TELECOMMUNICATIONS, volumen 2, p. 1177–1183, Papeete, French Polynesia, February 2003.

[29] PAPOULIS, A.. **Probability, Random Variables and Stochastic Processes**. McGraw-Hill, New York, 1965.

[30] SHANNON, C. E.. **Coding theorems for a discrete source with a fidelity criterion**. IRE National Convention Record, 7:142–163, 1959.

[31] SHANNON, C. E.. **Communication in the presence of noise**. Proceedings of the IEEE, 86:447–457, 1998.

[32] SHUM, K. W.. **Permutation coding and MFSK modulation for frequency selective channel**. In: PROC. IEEE INT. SYMP. PERSONAL, INDOOR AND MOBILE RADIO COMMUNICATIONS, volumen 5, p. 2063–2066, Lisbon, Portugal, September 2002.

[33] SILVA, D.; FINAMORE, W. A.. **Vector permutation code design algorithm**. In: PROC. INT. SYMP. INFORMATION THEORY AND ITS APPLICATIONS, Parma, Italy, October 2004.

[34] SLEPIAN, D.. **Permutation modulation**. Proceedings of the IEEE, 53:228–236, 1965.

[35] TOWNES, S. A.; J. B. O'NEAL, J.. **Permutation codes for the Laplacian source**. IEEE Transactions on Information Theory, 30:553–559, 1984.

[36] UNGERBOECK, G.. **Channel coding with multilevel/phase signals**. IEEE Transactions on Information Theory, 28:55–67, 1982.

[37] VITERBO, E.. **Trellis decoding of permutation modulations**. In: PROC. IEEE INT. SYMP. INFORMATION THEORY, p. 393, Yokohama, Japan, June 2003.

[38] WOLF, J. K.. **Permutation codes, (d,k) codes and magnetic recording**. In: PROC. IEEE COLLOQUIUM IN SOUTH AMERICA, p. 59–61, Argentina, Brazil, Chile, Uruguay, September 1990.

[39] WOZENCRAFT, J.; JACOBS, I. M.. **Principles of Communication Engineering**. Wiley, New York, 1965.