AULA-5 (11/03/2006) APROXIMAÇÃO DE FUNÇÕES: O MÉTODO DOS QUADRADOS MÍNIMOS LINEARES (MQML)

RESUMO: *Nesta aula continuaremos estudando como aproximar funções reais usando o método dos quadrados mínimos lineares (QML). Abordadremos o caso em que os dados são de tipo discreto (dados tabelados) e também o caso contínuo. Mostraremos que a técnica de regressão linear, ferramenta usual em estatística, é um caso especial do método dos QML. Apresentaremos alguns exemplos numéricos.*

O Método QML: Dado um vetor f num espaço vetorial V, o método QML consiste em calcular a *projeção ortogonal de* f num subespaço de dimensão finita $V_n \subset V$, e escolher a projeção como *aproximação* para f. No caso de dados discretos $V = \mathbb{R}^m$, e a projeção usa o produto interno definido por:

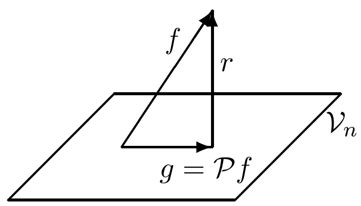
$$< x, y >= x^T y \equiv x_1 y_1 + x_2 y_2 + \cdots + x_m y_m. \tag{1}$$

Já para o caso de dados contínuos, o espaço escolhido é $C[a, b]$: o espaço de funções contínuas. Para $f,\ g \in C[a, b]$, o produto interno $< f, g >$ é difinido por:

$$< f, g >= \int_a^b f(x) g(x) dx. \tag{2}$$

Independentemente do tipo de dados, para aproximar a função f pelo método QML, o vetor g procurado deve satisfazer a condição de que o resíduo $r = f - g$ seja perpendicular a $\mathcal{V}_n$ (veja Fig.)

Critério: $r \perp \mathcal{V}_n$



Na prática, se $\phi_1, \phi_2, \ldots, \phi_n$ é uma base de $\mathcal{V}_n$, o vetor procurado $g = \mathcal{P}f$ deve ser escrito como

$$g = c_1\phi_1 + c_2\phi_2 + \cdots + c_n\phi_n$$

e os coeficientes c_i $(i = 1, \ldots, n)$ devem satisfazer a condição $r = (f - g) \perp \phi_j$. Aplicando o critério, segue que o vetor de coeficientes c_j's é solução do sistema de equações lineares

$$Ac = b, \quad [A]_{i,j} =< \phi_i, \phi_j >, \quad b_i =< \phi_i, f >, \quad i, j = 1, n. \tag{3}$$

O vetor g assim calculado é chamado *o melhor ajuste* de f no sentido dos quadrados minimos, ou também, *a melhor aproximação* de f em $\mathcal{V}_n$. As equações em (3) são chamadas *Equações normais.*

Regressão Linear: Se $y_i \approx g(x_i)$, $i = 1, \ldots, m$, e se assumimos que g é uma função do tipo $g(x) = c_0 + c_1 x$, onde c_0, c_1 são constantes desconhecidas, uma maneira de se estimar essas constantes é através do método QML. Para tal, seja $\mathbf{y} = [y_1, \ldots, y_m]^T$ o vetor de informações, $\mathbf{g} = [g(x_1), \ldots, g(x_m)]^T$ o vetor de valores da função g (desconhecida), e $\mathbf{r} = [r_1, \ldots, r_m]^T$ com $r_i = y_i - g(x_i)$ o vetor resíduo. Então, já que $\mathbf{g} = c_1\phi_1 + c_2\phi_2$, onde $\phi_1 = [1, 1, \ldots, 1]^T \in \mathbb{R}^m$ e $\phi_2 = [x_1, \ldots, x_m]^T$, a projeção de $\mathbf{y}$ no subespaço gerado por ϕ_1, ϕ_2 é calculada resolvendo-se o sistema de equações normais (3), onde

$$A = \begin{bmatrix} < \phi_1, \phi_1 > & < \phi_1, \phi_2 > \\ < \phi_2, \phi_1 > & < \phi_2, \phi_2 > \end{bmatrix} = \begin{bmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{bmatrix}, \ b = \begin{bmatrix} < \phi_1, \mathbf{y} > \\ < \phi_2, \mathbf{y} > \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \end{bmatrix}.$$

Exercícios

1. Escreva as equações normais para determinar o melhor ajuste quadrático $g(x) = c_0 + c_1 x + c_2 x^2$ de um conjunto de informações (x_i, y_i), $i = 1, \ldots, m$.

2. Determine um polinômio $p(x) = c_0 + c_1 x$ que melhor aproxima $f(x) = e^{-x}$ no intervalo [0,1]. Ilustre os resultados num gráfico. Qual é o vetor resíduo ?

Fpolis, 11/03/2006

Fermín S. V. Bazán