CCI-22

Matemática Computacional

Carlos Henrique Q. Forster

CCI-22

Ajuste de Curvas

Notas complementares

Ajuste de Curvas

- Aplica-se nos seguintes casos:
  - Extrapolação: valores fora do intervalo tabelado
  - Valores com erros (provenientes de observações)
- Consiste em:
  - Determinar parâmetros que definem uma função (curva) média para representar os dados
- Exemplo: Regressõ linear
  - Encontrar os parâmetros de uma reta que aproxime os dados observados de um experimento
  - Modelo para b medido em função de t:

$$b = c + d \cdot t$$

Regressão Linear

- Exemplos de fenômenos modelados linearmente:
  - Movimento uniforme -- b: posição, t: tempo, d: velocidade, c: posição inicial
  - Custo de produzir peças – b: custo total, t: número de peças, c: custo fixo para iniciar a produção, d: custo dependente do número de peças
  - Medida de deformação – b: deformação, t: carga, c: posição de repouso, d: coeficiente de deformação

Regressão Linear

- Várias medidas (b_i,t_i):
  - Se tivesse apenas 2 medidas, existe uma única reta: caindo no caso da interpolação (curva que passa por todos os pontos)
  - Havendo mais pontos, o sistema de equações resultante é superdeterminado. Se não há solução, o sistema pode ser utilizado para encontrar valores para os parâmetros que se aproxime o máximo de uma eventual solução assumindo que existe erro (estatisticamente distribuído) dos dados de entrada.

Regressão Linear

- Para m medidas (observações):

$$\begin{cases} c + d \cdot t_1 = b_1 \\ c + d \cdot t_2 = b_2 \\ \dots \\ c + d \cdot t_m = b_m \end{cases}$$

Forma matricial:

$$\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \\ 1 & t_m \end{bmatrix} \cdot \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Sistema Ax=b superdeterminado.

Pode não ter solução, depende não só de A, mas também de b.

Sistemas Superdeterminados

- Encontramos uma "solução" que minimize os resíduos
  - Os resíduos são obtidos como um vetor
  - Deve-se escolher um critério para minimizar: alguma norma vetorial

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \\ a_{31}x_1 + a_{32}x_2 = b_3 \end{cases} \longrightarrow \begin{cases} r_1 = a_{11}x_1 + a_{12}x_2 - b_1 \\ r_2 = a_{21}x_1 + a_{22}x_2 - b_2 \\ r_3 = a_{31}x_1 + a_{32}x_2 - b_3 \end{cases}$$

Critérios sobre os resíduos

- Norma L-1 (Manhatan)

$$E = |r_1| + |r_2| + |r_3|$$

- Norma L-∞ (Chebyshev)

$$E = \max\{|r_1|, |r_2|, |r_3|\}$$

- Norma L-2 (Euclides)

$$E = \sqrt{r_1^2 + r_2^2 + r_3^2}$$

Método dos Mínimos Quadrados

- Buscamos os valores dos parâmetros que minimizam o quadrado da norma euclidiana dos resíduos:
  - A função é diferenciável em relação ao resíduos.
  - O termo é quadrático, sendo a derivada linear, obtemos um sistema linear a resolver.

$$E = r_1^2 + r_2^2 + r_3^2$$

$$\nabla_{(x_1,x_2)} E = \mathbf{0} \longrightarrow \begin{cases} \dfrac{\partial E}{\partial x_1} = 0 \\ \dfrac{\partial E}{\partial x_2} = 0 \end{cases}$$

Método dos Mínimos quadrados

$$E = (a_{11}x_1 + a_{12}x_2 - b_1)^2 + (a_{21}x_1 + a_{22}x_2 - b_2)^2 + (a_{31}x_1 + a_{32}x_2 - b_3)^2$$

$$\frac{\partial E}{\partial x_1} = \begin{array}{l} 2(a_{11}x_1 + a_{12}x_2 - b_1)a_{11} \\ +2(a_{21}x_1 + a_{22}x_2 - b_2)a_{21} \\ +2(a_{31}x_1 + a_{32}x_2 - b_3)a_{31} \end{array} = 0$$

$$\frac{\partial E}{\partial x_2} = \begin{array}{l} 2(a_{11}x_1 + a_{12}x_2 - b_1)a_{12} \\ +2(a_{21}x_1 + a_{22}x_2 - b_2)a_{22} \\ +2(a_{31}x_1 + a_{32}x_2 - b_3)a_{32} \end{array} = 0$$

Reagrupando...

Método dos Mínimos Quadrados

- Obtemos o sistema linear, chamado "sistema de equações normais", cuja solução é a "solução de mínimos quadrados" do sistema superdeterminado.

$$(a_{11}a_{11} + a_{21}a_{21} + a_{31}a_{31})x_1 + (a_{12}a_{11} + a_{22}a_{21} + a_{32}a_{31})x_2 = a_{11}b_1 + a_{21}b_2 + a_{31}b_3$$
$$(a_{12}a_{11} + a_{22}a_{21} + a_{32}a_{31})x_1 + (a_{12}a_{12} + a_{22}a_{22} + a_{32}a_{32})x_2 = a_{12}b_1 + a_{22}b_2 + a_{32}b_3$$

Mínimos quadrados (forma matricial)

- A solução de mínimos quadrados do sistema superdeterminado
  Ax=b
  é equivalente à solução do sistema:
  $A^TAx=A^Tb$.
- Observe que A^TA é uma matriz quadrada e simétrica, levando a uma solução pelo método de Cholesky.
- Como $x=(A^TA)^{-1}A^Tb$, a matriz $(A^TA)^{-1}A^T$ é chamada pseudo-inversa da matriz A.

Regressão Linear

- Voltando ao caso da regressão linear:

$$\begin{bmatrix}1 & t_1\\1 & t_2\\\vdots & \\1 & t_m\end{bmatrix}\cdot\begin{bmatrix}c\\d\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}$$

Multiplicando ambos lados pela transposta:

$$\begin{bmatrix}1 & 1 & \cdots & 1\\t_1 & t_2 & \cdots & t_m\end{bmatrix}\begin{bmatrix}1 & t_1\\1 & t_2\\\vdots & \\1 & t_m\end{bmatrix}\cdot\begin{bmatrix}c\\d\end{bmatrix}=\begin{bmatrix}1 & 1 & \cdots & 1\\t_1 & t_2 & \cdots & t_m\end{bmatrix}\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}$$

Regressão Linear

$$\begin{bmatrix}1 & 1 & \cdots & 1\\t_1 & t_2 & \cdots & t_m\end{bmatrix}\begin{bmatrix}1 & t_1\\1 & t_2\\\vdots & \\1 & t_m\end{bmatrix}\cdot=\begin{bmatrix}m & \sum_{i=1}^m t_i\\\sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2\end{bmatrix}$$

$$\begin{bmatrix}1 & 1 & \cdots & 1\\t_1 & t_2 & \cdots & t_m\end{bmatrix}\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}=\begin{bmatrix}\sum_{i=1}^m b_i\\\sum_{i=1}^m t_i b_i\end{bmatrix}$$

$$\begin{bmatrix}m & \sum_{i=1}^m t_i\\\sum_{i=1}^m t_i & \sum_{i=1}^m t_i^2\end{bmatrix}=\begin{bmatrix}\sum_{i=1}^m b_i\\\sum_{i=1}^m t_i b_i\end{bmatrix}$$

Ajuste de reta

Exemplo

t	b
-1	1
1	1
2	3

$$\begin{cases}c-d=1\\c+d=1\\c+2d=3\end{cases}\longrightarrow\begin{bmatrix}1 & -1\\1 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}c\\d\end{bmatrix}=\begin{bmatrix}1\\1\\3\end{bmatrix}$$

$$A^TAx=A^Tb$$

$$\begin{bmatrix}3 & 2\\2 & 6\end{bmatrix}\begin{bmatrix}c\\d\end{bmatrix}=\begin{bmatrix}5\\6\end{bmatrix}$$

$$c=\frac{9}{7},\quad d=\frac{4}{7}$$

Melhor reta é

$$b=\frac{9}{7}+\frac{4}{7}t$$

Ajuste de parábola

$$b=c+d\cdot t+e\cdot t^2$$

$$\begin{bmatrix}1 & t_1 & t_1^2\\1 & t_2 & t_2^2\\\vdots & & \\1 & t_m & t_m^2\end{bmatrix}\cdot\begin{bmatrix}c\\d\\e\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}$$

$$\begin{bmatrix}m & \sum t & \sum t^2\\\sum t & \sum t^2 & \sum t^3\\\sum t^2 & \sum t^3 & \sum t^4\end{bmatrix}\begin{bmatrix}c\\d\\e\end{bmatrix}=\begin{bmatrix}\sum b\\\sum bt\\\sum bt^2\end{bmatrix}$$

Mal-condicionamento da matriz de ajuste

- Supondo que o conjunto de amostra foi obtido para todos pontos no intervalo de 0 a 1, a matriz de regressão seria a seguinte:

$$\begin{bmatrix} \int_0^1 dx & \int_0^1 x\,dx & \int_0^1 x^2 dx \\ \int_0^1 x\,dx & \int_0^1 x^2 dx & \int_0^1 x^3 dx \\ \int_0^1 x^2 dx & \int_0^1 x^3 dx & \int_0^1 x^4 dx \end{bmatrix} = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}$$

Que é a famosa matriz de Hilbert, cujo número de condição cresce muito rápido com sua dimensão.

Aproximação por uma base de funções genéricas

- Definimos:
  - Função f(x) é a que queremos aproximar e da qual temos apenas observações amostradas nos pontos: x_i, 1<=i<=m
  - Uma sequência de funções $g_j(x)$, 1<=j<=n, que forma uma base defunções, ou seja, são linearmente independentes nos pontos x_i.
  - A função aproximadora φ(x) é uma combinação linear das funções da base, cujos coeficientes a_j devem ser obtidos

$$\varphi(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \cdots + \alpha_n g_n(x)$$

Ajuste por base de funções

- Os coeficientes são determinados de forma a minimizar a norma dos resíduos nos pontos amostrados

$$E = \| f(x) - \varphi(x) \|$$

Essa norma pode se basear no critério dos mínimos quadrados aplicado nos pontos observados

$$E(\alpha_1, \alpha_2, \cdots \alpha_n) = \sum_{i=1}^{m} \left(f(x_i) - \varphi(x_i) \right)^2$$

$$E(\alpha_1, \alpha_2, \cdots \alpha_n) = \sum_{i=1}^{m} \left(f(x_i) - \sum_{j=1}^{n} \alpha_j g_j(x_i) \right)^2$$

- Calculam-se as derivadas parciais em função de cada parâmetro aj e iguala-se a zero para obter o mínimo E.

$$\forall j, 1 \le j \le n \left\{ \frac{\partial E}{\partial \alpha_j} = 2 \sum_{i=1}^{m} \left[f(x_i) - \sum_{k=1}^{n} \alpha_k g_k(x_i) \right] \cdot \left[- g_j(x_i) \right] = 0 \right.$$

Antes de abrir o sistema de equações e isolar as incógnitas (os parâmetros a_j), vamos definir uma notação para os somatórios dos produtos das funções nos pontos observados...

- Definimos o produto escalar de duas funções em relação a um conjunto de pontos observados:

$$\langle g_1, g_2 \rangle = \sum_{i=1}^{m} g_1(x_i) \cdot g_2(x_i)$$

Decorre uma definição de norma (ao quadrado):

$$\langle g_1, g_1 \rangle = \sum_{i=1}^{m} [g_1(x_i)]^2$$

- Agora, expandindo o sistema de equações normais, obtemos:

$$\begin{bmatrix} \langle g_1, g_1 \rangle & \langle g_1, g_2 \rangle & \cdots & \langle g_1, g_n \rangle \\ \langle g_2, g_1 \rangle & \langle g_2, g_2 \rangle & \cdots & \langle g_2, g_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle g_n, g_1 \rangle & \langle g_n, g_2 \rangle & \cdots & \langle g_n, g_n \rangle \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \langle g_1, f \rangle \\ \langle g_2, f \rangle \\ \vdots \\ \langle g_n, f \rangle \end{bmatrix}$$

Observe a simetria do sistema obtido, dado que

$$\langle g_1, g_2 \rangle = \langle g_2, g_1 \rangle$$

Assim, uma vez calculados os produtos internos, pode-se resolver o sistema por Cholesky e obter os coeficientes que definem a função φ(x).

Exemplos de bases de funções

$$\begin{cases} g_1(x) = 1 \\ g_2(x) = x \\ g_3(x) = x^2 \end{cases}$$

$$\begin{cases} g_1(x) = 1 \\ g_2(x) = x \\ g_3(x) = \frac{3}{2}x^2 - \frac{1}{2} \end{cases}$$

Legendre

$$\begin{cases} g_1(x) = 1 \\ g_2(x) = \cos(x) \\ g_3(x) = \sin(x) \\ g_4(x) = \cos(2x) \\ g_5(x) = \sin(2x) \\ g_6(x) = \cos(3x) \\ g_7(x) = \sin(3x) \end{cases}$$

Trigonométrica (Fourier)

Caso contínuo

- É possível estabelecer um caso contínuo, em que se conheça a função f(x) e se deseja obter a aproximação φ(x) mais aderente em uma dado intervalo.
- O critério de mínimos quadrados é então:

$$E = \int_I [f(x) - \varphi(x)]^2 dx$$

E o produto interno das funções é definido por

$$\langle g_1, g_2 \rangle = \int_I g_1(x) \cdot g_2(x) dx$$

Base ortogonal

- Uma base ortogonal é definida pela propriedade:

$$\langle g_i, g_j \rangle = \begin{cases} 0, \text{se } i \neq j \\ \neq 0, \text{se } i = j \end{cases} \quad \text{(=1 se ortonormal)}$$

Nessas condições, o sistema a ser resolvido passa a ser diagonal.

$$\begin{bmatrix} \langle g_1, g_1 \rangle & 0 & \cdots & 0 \\ 0 & \langle g_2, g_2 \rangle & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \langle g_n, g_n \rangle \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \langle g_1, f \rangle \\ \langle g_2, f \rangle \\ \vdots \\ \langle g_n, f \rangle \end{bmatrix}$$

Base ortogonal

- A ortogonalidade vai depender do conjunto de pontos observados ou do intervalo de integração
- A base trigonométrica será ortogonal para integração de 0 a 2π.

$$\begin{cases} g_1(x) = 1 \\ g_2(x) = x \\ g_3(x) = \frac{3}{2}x^2 - \frac{1}{2} \end{cases}$$

- A base de Legendre será ortogonal para integração no intervalo de -1 a 1

Base ortogonal

- Uma base ortogonal pode ser obtida através dos passos de Gram-Schmidt.
- Sejam as funções a(x), b(x) e c(x) uma base (lin. indep. Portanto), determinamos uma base ortogonal a'(x),b'(x),c'(x):

$$a'(x) = a(x)$$

$$b'(x) = b(x) - \frac{\langle a', b \rangle}{\langle a', a' \rangle} a'(x)$$

$$c'(x) = c(x) - \frac{\langle c, b' \rangle}{\langle b', b' \rangle} b'(x) - \frac{\langle c, a' \rangle}{\langle a', a' \rangle} a'(x)$$

Decomposição QR

- Vamos aplicar os passos da ortogonalização de Gram-Schmidt nas colunas de uma matriz.

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{onde} \quad A = \begin{bmatrix} \vdots & \vdots & \vdots \\ a & b & c \\ \vdots & \vdots & \vdots \end{bmatrix}$$

$$b' = b - \frac{a^T b}{a^T a} a = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix}$$

$$c' = c - \frac{b'^T c}{b'^T b'} b' - \frac{a'^T c}{a'^T a'} a' = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{\frac{1}{2}} \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Decomposição QR

- Observe que a matriz é obtida a partir de uma transformação triangular da matriz original. Normalizamos as colunas...

$$\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 1 & -\frac{1}{2} & 0 \end{bmatrix} \longrightarrow \quad Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}$$

A matriz Q é ortonormal, de forma que $Q^TQ=I$

Podemos obter a matriz R através de $R=Q^TA$

A decomposição A=QR é obtida, com Q ortonormal e R triangular superior.