MMQ e Projeção Ortogonal

Dados os vetores $v^0 \in \mathbb{R}^5$; $v^1 \in \mathbb{R}^5$ e $u \in \mathbb{R}^5$

$$v^0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}; \quad v^1 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}; \quad u = \begin{pmatrix} 0.5 \\ 11 \\ 19.7 \\ 33.7 \\ 40.3 \end{pmatrix};$$

Considere os subespaços vetoriais $V \subset \mathbb{R}^5$ definido pelas combinações lineares de v^0 e v^1 e $V_\perp \subset \mathbb{R}^5$ o complemento ortogonal de V. A seguir vamos obter a decomposição do vetor u como uma soma $u_V + u_{V_\perp}$, onde u_V é a projeção ortogonal de u no subespaço V e $u_{V_\perp}$ é a projeção ortogonal de u no subespaço $V_\perp$.

Como u_V é um vetor no subespaço V temos que:

$$u_V = a_0 v^0 + a_1 v^1$$

Assim, a decomposição de u como uma soma $u = u_V + u_{V_\perp}$ fica definida a partir dos valores de a_0 e a_1. Equações envolvendo os parâmetros a_0 e a_1 podem ser obtidas a partir da ortogonalidade de $u_{V_\perp}$ e v^0 e da ortogonalidade de $u_{V_\perp}$ e v^1.

Considerando em $\mathbb{R}^5$ uma geometria Euclideana ou, de forma equivalente, a definição de angulo obtida a partir do produto interno

$$\langle u \mid w \rangle = \sum_{i=1}^{5} u_i w_i$$

temos

$$\langle v^0 \mid u \rangle = \langle v^0 \mid u_V + u_{V_\perp} \rangle$$
$$\langle v^1 \mid u \rangle = \langle v^1 \mid u_V + u_{V_\perp} \rangle$$

Como v^0 e v^1 são ortognais a $u_{V_\perp}$,

$$\langle v^0 \mid u_{V_\perp} \rangle = 0$$

e

$$\langle v^1 \mid u_{V_\perp} \rangle = 0$$

temos então:

$$\langle v^0 \mid u \rangle = \langle v^0 \mid u_V \rangle + \cancel{\langle v^0 \mid u_{V_\perp} \rangle} = \langle v^0 \mid u_V \rangle$$

$$\langle v^1 \mid u \rangle = \langle v^1 \mid u_V \rangle + \cancel{\langle v^1 \mid u_{V_\perp} \rangle} = \langle v^1 \mid u_V \rangle$$

de onde obtemos o seguinte sistema de equações lineares:

$$\langle v^0 \mid u_V \rangle = \langle v^0 \mid a_0 v^0 + a_1 v^1 \rangle = \langle v^0 \mid u \rangle$$

$$\langle v^1 \mid u_V \rangle = \langle v^1 \mid a_0 v^0 + a_1 v^1 \rangle = \langle v^1 \mid u \rangle$$

$$\langle v^0 \mid v^0 \rangle a_0 + \langle v^0 \mid v^1 \rangle a_1 = \langle v^0 \mid u \rangle$$

$$\langle v^1 \mid v^0 \rangle a_0 + \langle v^1 \mid v^1 \rangle a_1 = \langle v^1 \mid u \rangle$$

ou na forma matricial

$$\begin{pmatrix} \langle v^0 \mid v^0 \rangle & \langle v^0 \mid v^1 \rangle \\ \langle v^1 \mid v^0 \rangle & \langle v^1 \mid v^1 \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \langle v^0 \mid u \rangle \\ \langle v^1 \mid u \rangle \end{pmatrix}$$

Calculando os produtos internos:

$\langle v^0 \mid v^0 \rangle$, $\langle v^0 \mid v^1 \rangle$, $\langle v^1 \mid v^1 \rangle$, $\langle v^0 \mid u \rangle$ e $\langle v^1 \mid u \rangle$;

$$\langle v^0 \mid v^0 \rangle = 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 = 5$$

$$\langle v^0 \mid v^1 \rangle = 1 \cdot 0 + 1 \cdot 1 + 1 \cdot 2 + 1 \cdot 3 + 1 \cdot 4 = 10$$

$$\langle v^1 \mid v^1 \rangle = 0 \cdot 0 + 1 \cdot 1 + 2 \cdot 2 + 3 \cdot 3 + 4 \cdot 4 = 30$$

$$\langle v^0 \mid u \rangle = 1 \cdot 0.5 + 1 \cdot 11 + 1 \cdot 19.7 + 1 \cdot 33.7 + 1 \cdot 40.3 = 105.2$$

$$\langle v^1 \mid u \rangle = 0 \cdot 0.5 + 1 \cdot 11 + 2 \cdot 19.7 + 3 \cdot 33.7 + 4 \cdot 40.3 = 312.7$$

De forma que para obter os coeficientes a_0 e a_1 que definem u_v, devemos resolver:

$$\begin{pmatrix} 5 & 10 \\ 10 & 30 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 105.2 \\ 312.7 \end{pmatrix}$$

$$a_0 = \frac{105.2 \cdot 30 - 10 \cdot 312.7}{5 \cdot 30 - 10 \cdot 10} = 0.58$$

$$a_1 = \frac{5 \cdot 312.7 - 105.2 \cdot 10}{5 \cdot 30 - 10 \cdot 10} = 10.23$$

$$u_v = 0.58v^0 + 10.23v^1$$

Uma outra questão envolvendo os vetores v^0, v^1 e u seria determinar o vetor em V mais próximo de u, considerando a noção de distância definida a partir do produto interno Euclideano.

Dados u e w vetores em $\mathbb{R}^n$ e um produto interno $\langle \cdot \mid \cdot \rangle$, a distância entre u e v é definida por:

$$dist(u, w) = \sqrt{\langle u - w \mid u - w \rangle}$$

Para cada escolha de coeficientes a_0 e a_1, o quadrado da distância entre o vetor $v = a_0 v^0 + a_1 v^1 \in V$ e o vetor u é dada por:

$$dist^2(v, u) = \langle \left(a_0 v^0 + a_1 v^1\right) - u \mid \left(a_0 v^0 + a_1 v^1\right) - u \rangle = D(a_0, a_1)$$

Assim, para determinar o vetor em V mais próximo de u, devemos encontrar os valores de a_0 e a_1 para os quais $D(a_0, a_1)$ é mínimo.

Portanto devemos resolver:

$$\frac{\partial D}{\partial a_0}(a_0, a_1) = 0$$

$$\frac{\partial D}{\partial a_1}(a_0, a_1) = 0$$

$$\frac{\partial D}{\partial a_i}(a_0, a_1) = \frac{\partial}{\partial a_i}\langle \left(a_0 v^0 + a_1 v^1\right) - u \mid \left(a_0 v^0 + a_1 v^1\right) - u \rangle =$$

$$\frac{\partial}{\partial a_i}\left\{\langle v^0 \mid v^0\rangle a_0^2 + \langle v^0 \mid v^1\rangle a_0 a_1 + \langle v^1 \mid v^0\rangle a_1 a_0 + \langle v^1 \mid v^1\rangle a_1^2\right.$$

$$\left. -2\langle v^0 \mid u\rangle a_0 - 2\langle v^1 \mid u\rangle a_1\right\}$$

$$\frac{\partial D}{\partial a_0}(a_0, a_1) = 0 \iff$$
$$\langle v^0 \mid v^0 \rangle a_0 + \langle v^0 \mid v^1 \rangle a_1 - \langle v^0 \mid u \rangle = 0$$

$$\frac{\partial D}{\partial a_1}(a_0, a_1) = 0 \iff$$
$$\langle v^1 \mid v^0 \rangle a_0 + \langle v^1 \mid v^1 \rangle a_1 - \langle v^1 \mid u \rangle = 0$$

Assim os coeficientes a_0 e a_1 que definem o vetor $v = a_0 v^0 + a_1 v^1 \in V$ mais próximo de u são obtidos resolvendo:

$$\begin{pmatrix} \langle v^0 \mid v^0 \rangle & \langle v^0 \mid v^1 \rangle \\ \langle v^1 \mid v^0 \rangle & \langle v^1 \mid v^1 \rangle \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \langle v^0 \mid u \rangle \\ \langle v^1 \mid u \rangle \end{pmatrix}$$

de onde obtemos:

$$v = 0.58v^0 + 10.23v^1$$

Estas observações no contexto de Algebra Linear nos leva às seguintes interpretações do Método dos Mínimos Quadrados:

Dados:

1. **Modelo** - Uma formulação explicita da dependência de y em função das variáveis $\mathbf{x} = (x_1, \ldots, x_\ell)$, expressa em termos de um conjunto de parâmetros $\{a_0, a_1, \ldots, a_m\}$; $a_k \in \mathbb{R}$ e, uma família de funções $\{g_0(\mathbf{x}), g_1(\mathbf{x}), \ldots, g_m(\mathbf{x})\}$; $g_k() : \mathbb{R}^\ell \to \mathbb{R}$.

$$y = f(\mathbf{x}) = a_0 g_0(\mathbf{x}) + a_1 g_1(\mathbf{x}) + \ldots + a_m g_m(\mathbf{x})$$

2. **Resultados Empíricos** (*caso discreto*)
Um conjunto de dados obtidos pela observação de valores da variável dependente y para diferentes valores das variáveis independentes $\mathbf{x}$,
$\left\{(\mathbf{x}^i, y_i)\right\}_{i=1,\ldots,n}$; $\mathbf{x}^i \in \mathbb{R}^\ell$; $y_i \in \mathbb{R}$.

Considere vetores $\left\{v^j\right\}_{j=0,\dots m}$; $v^j \in \mathbb{R}^n$ definidos por:

$$v^j = \begin{pmatrix} v_1^j \\ v_2^j \\ \vdots \\ v_n^j \end{pmatrix} = \begin{pmatrix} g_j(\mathbf{x}^1) \\ g_j(\mathbf{x}^2) \\ \vdots \\ g_j(\mathbf{x}^n) \end{pmatrix}$$

e o vetor $u \in \mathbb{R}^n$ definido por:

$$u = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Assim obter o *melhor ajuste* no sentido do Método dos Mínimos Quadrados é equivalente a obter a *projeção ortogonal* de u no subespaço V definido pelas combinações lineares dos vetores $\{v^j\}_{j=0,\ldots m}$ ou determinar o vetor $v \in V$ mais próximo de u.

A interpretação do Método dos Mínimos Quadrados tanto como uma problema de projeção ortogonal ou determinar o vetor $v \in V$ mais próximo de u permite generalizar o Método no sentido de que no contexto das ideias de Algebra Linear podemos considerar as questões envolvidas em um nível abstrato sem a necessidade de especificar o espaço vetorial ou produto interno envolvidos.