MATRIZES DE PROJEÇÃO

SEVERINO TOSCANO DO REGO MELO
USP – MAT 2458 – TURMA 20 – 2014

Dada uma base de um subespço W de $\mathbb{R}^n$, queremos encontrar a matriz P da projeção ortogonal de $\mathbb{R}^n$ sobre W. Nosso objetivo é demonstrar a fórmula (2), que expressa P como função da matriz A que tem nas colunas as coordenadas (relativas à base canônica) dos vetores da base dada.

Estas notas são baseadas em [1, Seção 6.4].

1. Notação e Resultados Preliminares

Denotemos por $\operatorname{Im} A$ o espaço-coluna da matriz A (em outras palavras, se A é $n \times k$, $\operatorname{Im} A$ é a imagem da transformação linear $T_A : \mathbb{R}^k \to \mathbb{R}^n$ cuja matriz relativa à base canônica é A); e por $\operatorname{Ker} A$ o espaço das soluções do sistema linear homogêneo cuja matriz de coeficientes é A (em outras palavras, $\operatorname{Ker} A$ é o núcleo de T_A).

Dado um subespaço vetorial W de $\mathbb{R}^n$, denotamos por $W^\perp$ o subespaço que consiste dos vetores de $\mathbb{R}^n$ que são perpendiculares (relativamente ao produto interno canônico) a todos os elementos de W.

Vamos usar os seguintes resultados:

(1) As colunas de uma matriz A são linearmente independentes se, e somente se, $\operatorname{Ker} A$ consiste apenas do vetor nulo. Essas condições são também equivalentes a afirmar que, sempre que $\mathbf{y} \in \mathbb{R}^n$ é tal que o sistema linear $A\mathbf{x} = \mathbf{y}$ tem solução $\mathbf{x} \in \mathbb{R}^k$, então essa solução é única.
(2) Se o espaço-coluna de uma matriz $k \times k$ é igual a $\mathbb{R}^k$, então a matriz é inversível.
(3) Dado um subespaço W de $\mathbb{R}^n$, a projeção ortogonal de um vetor $\mathbf{b} \in \mathbb{R}^n$ sobre W é o único vetor $P\mathbf{b}$ que satisfaz $P\mathbf{b} \in W$ e $(\mathbf{b} - P\mathbf{b}) \in W^\perp$.
(4) A igualdade $\operatorname{Ker} A^t = (\operatorname{Im} A)^\perp$ vale para toda matriz A (se A é $n \times k$, a transposta A^t é uma matriz $k \times n$ e, portanto, $\operatorname{Ker} A^t$ e $(\operatorname{Im} A)^\perp$ são subespaços de $\mathbb{R}^n$).

2. Solução do problema

Vamos usar as quatro proposições listadas acima, sem alertar que as estamos usando. Fica por conta do leitor perceber em qual momento cada uma delas estará sendo usada.

Dado qualquer $\mathbf{b} \in \mathbb{R}^n$, o sistema linear $A\mathbf{x} = P\mathbf{b}$ tem solução $\mathbf{x} \in \mathbb{R}^k$, pois $P\mathbf{b} \in W = \operatorname{Im} A$. Além disso, a solução é única, pois as colunas de A são linearmente independentes. Esse único $\mathbf{x}$, que satisfaz $A\mathbf{x} = P\mathbf{b}$, obviamente satisfaz também $A^tA\mathbf{x} = A^tP\mathbf{b}$.

O pulo do gato vem agora: $A^t(\mathbf{b} - P\mathbf{b}) = 0$, pois $(\mathbf{b} - P\mathbf{b}) \in (\operatorname{Im} A)^{\perp} = \operatorname{Ker} A^t$, pois $P\mathbf{b}$ é a projeção ortogonal de $\mathbf{b}$ sobre $W = \operatorname{Im} A$. Daí, podemos escrever:

$$A^tA\mathbf{x} = A^tP\mathbf{b} = A^tP\mathbf{b} + A^t(\mathbf{b} - P\mathbf{b}) = A^t\mathbf{b}. \tag{1}$$

A matriz A^tA é quadrada, de tamanho $k \times k$. O próximo passo é provar que ela é inversível.

Todo vetor pertencente a $\operatorname{Im} A^tA$ pertence também a $\operatorname{Im} A^t$ (pois $(A^tA)\mathbf{x} = A^t(A\mathbf{x})$, para todo $\mathbf{x} \in \mathbb{R}^k$). Vamos provar agora que vale a recíproca, isto é: se um vetor $\mathbf{y}$ pertence a $\operatorname{Im} A^t$, então ele pertence também a $\operatorname{Im} A^tA$. De fato, se $\mathbf{y} = A^t\mathbf{x}$, para algum $\mathbf{x} \in \mathbb{R}^n$, então, chamando de $\tilde{\mathbf{x}}$ a projeção ortogonal de $\mathbf{x}$ sobre $\operatorname{Im} A$, temos $(\tilde{\mathbf{x}} - \mathbf{x}) \in (\operatorname{Im} A)^{\perp}$ e, portanto:

$$\mathbf{y} = A^t\mathbf{x} = A^t\mathbf{x} + A^t(\tilde{\mathbf{x}} - \mathbf{x}) = A^t\tilde{\mathbf{x}}.$$

Como $\tilde{\mathbf{x}} \in \operatorname{Im} A$, existe $\mathbf{z} \in \mathbb{R}^k$ tal que $\tilde{\mathbf{x}} = A\mathbf{z}$. Daí, $\mathbf{y} = A^tA\mathbf{z}$, o que mostra que $\mathbf{y} \in \operatorname{Im} A^tA\mathbf{z}$.

Provamos, portanto, que $\operatorname{Im} A^t = \operatorname{Im} A^tA$.

As colunas de A são linearmente independentes, logo,

$$\{0\} = \operatorname{Ker} A = (\operatorname{Im} A^t)^{\perp}, \quad \text{daí}, \quad \mathbb{R}^k = \operatorname{Im} A^t = \operatorname{Im} A^tA,$$

o que mostra que A^tA é inversível.

Segue então de (1), que $\mathbf{x} = (A^tA)^{-1}A^t\mathbf{b}$. Mas $\mathbf{x}$ satisfaz $A\mathbf{x} = P\mathbf{b}$, logo temos que $P\mathbf{b} = A(A^tA)^{-1}A^t\mathbf{b}$. Isso vale para todo $\mathbf{b} \in \mathbb{R}^n$, o que demonstra a fórmula

$$P = A(A^tA)^{-1}A^t. \tag{2}$$

É interessante observar o que diz esta fórmula em dois casos extremos: quando $k = n$, então A e A^t são inversíveis e obtemos $P = I$; quando $k = 1$, obtemos a fórmula usual de projeção ortogonoal sobre um vetor não-nulo.

Verifica-se com um cálculo direto (é preciso usar que A^tA é autoadjunta e que, portanto, sua inversa também é) que a matriz P obtida em (2) satisfaz $P = P^2 = P^t$, como tinha de ser: pode-se demonstrar (com um argumento mais simples, embora um pouco mais abstrato do que o que usamos aqui) que uma matriz P é a matriz de uma projeção ortogonal se, se somente se, as igualdades $P = P^2 = P^t$ são satisfeitas.

Referências

[1] Howard Anton & Chris Rorres. Álgebra linear com aplicações. Bookman. Oitava Edição.