Experiência 03
Transformações Lineares

9 de março de 2015

Transformações em espaços vetoriais são funções que levam vetores de um espaço vetorial origem para outros vetores em um espaço vetorial destino. Recordando,

Definição 1. Uma *função* ou *transformação* f de um conjunto origem (também conhecido como *domínio* da função ou transformação) D para um conjunto destino (ou *contra-domínio* da transformação) E é uma relação entre elementos de D e E tal que, para todo elemento $x \in D$, existe um (e somente um) correspondente $y = f(x) \in E$.

Essa definição está colocada aqui para recordar que uma função é uma relação entre dois conjuntos (o domínio e o contra-domínio) que tem de satisfazer duas propriedades: 1) a função deve ser definida para *todos* os elementos do domínio; 2) o valor da função para cada elemento do domínio deve ser *único*. É importante lembrar dessas condições quando formos tratar de inversas de transformações lineares e matrizes, principalmente do que é possível "inverter" quando a matriz que não for quadrada.

Transformações *lineares* são funções que satisfazem propriedades de aditividade e proporcionalidade (homogeneidade):

Definição 2. Uma *transformação linear* de um espaço vetorial U para um espaço vetorial V (ambos sobre o mesmo conjunto de escalares $\mathbb{R}$) é uma função $f : U \to V$ satisfazendo

$$f(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y}),$$

para todos os vetores $\boldsymbol{x}, \boldsymbol{y} \in U$ e quaisquer $\alpha, \beta \in \mathbb{R}$.

Repare que a definição de transformação linear é bem restritiva: por exemplo, se T é uma transformação linear, então

$$T(\mathbf{0}_U) = T(\mathbf{0}_U + \mathbf{0}_U) = T(\mathbf{0}_U) + T(\mathbf{0}_U) \Rightarrow T(\mathbf{0}_U) = \mathbf{0}_V,$$

ou seja, uma transformação linear T de U em V sempre leva o vetor nulo de U, $\mathbf{0}_U$, para o vetor nulo de V, $\mathbf{0}_V$.

Se o conceito de transformação linear é assim "restritivo", por que então estudá-lo a fundo? As principais razões são:

1. A propriedade de linearidade e a estrutura de espaço vetorial permitem que o efeito de uma transformação linear seja bem compreendido, tanto do ponto de vista teórico, quanto, por meio de algoritmos eficientes, do ponto de vista prático. Isto significa que um problema real modelado por transformações lineares pode ser analisado de maneira eficiente e profunda usando as ferramentas de Álgebra Linear.

2. Muitas operações de interesse prático são lineares ou aproximadamente lineares, como por exemplo, a distorção observada em uma imagem feita por uma lente desfocada, a mudança das coordenadas de um corpo rígido devida a uma rotação, a derivada de uma função (e, sob certas condições, também a integral), um atraso em uma sequência, ou mais geralmente a transformação efetuada em uma sequência ou função por um filtro linear, a transformada de Fourier (de tempo discreto ou contínuo), e problemas envolvendo probabilidades, como o cálculo de estimativas de sinais Gaussianos segundo o critério dos mínimos quadrados.

3. Vários problemas cujas soluções ótimas não envolvem transformações lineares são resolvidos por meio sequências de transformações lineares devido aos algoritmos eficientes e ao estado avançado da teoria disponíveis para estas últimas; por exemplo, vários algoritmos iterativos para cálculo de zeros de funções ou para otimização, métodos para integração numérica de equações diferenciais, o cálculo de estimativas lineares para sinais não-Gaussianos, o filtro de Kalman estendido, etc.

Nas próximas experiências vamos ver algumas dessas aplicações. A maioria das outras será vista em cursos futuros, como Cálculo Numérico, Sistemas e Sinais, Controle, Processamento de Sinais, Probabilidade e Processos Estocásticos.

1 Matrizes e Transformações Lineares

Dada uma transformação linear $f : \mathbb{R}^n \to \mathbb{R}^m$ e bases[1] $\mathcal{B}_n = \{\boldsymbol{u}_1 \dots \boldsymbol{u}_n\}$ e $\mathcal{B}_m = \{\boldsymbol{v}_1 \dots \boldsymbol{v}_m\}$ para $\mathbb{R}^n$ e $\mathbb{R}^m$, podemos descrever a transformação linear por meio de uma matriz, da seguinte maneira. Para cada $\boldsymbol{x} \in \mathbb{R}^n$, temos

$$\boldsymbol{x} = \sum_{j=1}^{n} \alpha_j \boldsymbol{u}_j \Rightarrow \boldsymbol{y} = f(\boldsymbol{x}) = \sum_{j=1}^{n} \alpha_j f(\boldsymbol{u}_j).$$

Como $f(\boldsymbol{u}_j) \in \mathbb{R}^m$, e $\mathcal{B}_m$ é uma base de $\mathbb{R}^m$, vale $f(\boldsymbol{u}_j) = \sum_{i=1}^{m} \beta_{i,j} \boldsymbol{v}_i$. O resultado final é que $f(\boldsymbol{x})$ pode ser escrita em termos da base $\mathcal{B}_m$, como

$$f(\boldsymbol{x}) = \sum_{j=1}^{n} \alpha_j \left(\sum_{i=1}^{m} \beta_{i,j} \boldsymbol{v}_i \right) = \sum_{i=1}^{m} \boldsymbol{v}_i \left(\sum_{j=1}^{n} \beta_{i,j} \alpha_j \right).$$

[1] Recordando, se $\mathcal{B}_n = \{\boldsymbol{u}_1 \dots \boldsymbol{u}_n\}$ é uma base do $\mathbb{R}^n$, então $\{\boldsymbol{u}_1 \dots \boldsymbol{u}_n\}$ é linearmente independente e todo vetor $\boldsymbol{x} \in \mathbb{R}^n$ pode ser escrito como $\alpha_1 \boldsymbol{u}_1 + \alpha_2 \boldsymbol{u}_2 + \cdots + \alpha_n \boldsymbol{u}_n$, em que $\alpha_1, \dots, \alpha_n$ são constantes.

Agora, juntando os α_j e $\beta_{i,j}$ em forma de vetor e matriz,

$$\left(\boldsymbol{x}\right)_{\mathcal{B}_n} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \qquad\qquad \boldsymbol{A} = \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \dots & \beta_{1,n} \\ \beta_{2,1} & \beta_{2,2} & \dots & \beta_{2,n} \\ \vdots & & \ddots & \vdots \\ \beta_{m,1} & \beta_{m,2} & \dots & \beta_{m,n} \end{bmatrix},$$

$f(\boldsymbol{x})$, na base $\mathcal{B}_m$, se escreve como

$$\big(f(\boldsymbol{x})\big)_{\mathcal{B}_m} = \boldsymbol{A}.\left(\boldsymbol{x}\right)_{\mathcal{B}_n}.$$

Além disso, dadas $\mathcal{B}_n$, $\mathcal{B}_m$ e f, a matriz $\boldsymbol{A}$ é única. Vamos denotar os elementos da matriz $\boldsymbol{A}$ por $[\boldsymbol{A}]_{ij}$ ou por a_{ij}, o primeiro índice correspondendo a linhas e o segundo a colunas.

Repare que as definições habituais de soma e produto entre matrizes são definidas calculando-se a matriz que representa as transformações lineares $(f+g)(\boldsymbol{x}) \triangleq f(\boldsymbol{x})+g(\boldsymbol{x})$ e $(f\circ g)(\boldsymbol{x}) \triangleq f\big(g(\boldsymbol{x})\big)$.

Soma de matrizes: Sejam f e g duas transformações lineares de $\mathbb{R}^n$ para $\mathbb{R}^m$, e defina a *soma* de f e g por

$$(f+g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x}).$$

Então, dadas as bases $\mathcal{B}_n$ de $\mathbb{R}^n$ e $\mathcal{B}_m$ de $\mathbb{R}^m$, se a matriz correspondente a f for $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, e a matriz correspondente a g for $\boldsymbol{B} \in \mathbb{R}^{m\times n}$, então a matriz $\boldsymbol{C}$ correspondente a $f+g$ será

$$\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B},$$

em que o elemento c_{ij} de $\boldsymbol{C}$ é dado por $c_{ij} = a_{ij} + b_{ij}$.

Multiplicação de matrizes: Seja f uma transformação linear de $\mathbb{R}^p$ para $\mathbb{R}^m$, e g uma transformação linear de $\mathbb{R}^n$ para $\mathbb{R}^p$. Sejam também $\boldsymbol{A}$ e $\boldsymbol{B}$ as matrizes correspondentes a f e g, respectivamente, definidas a partir das bases $\mathcal{B}_n$, $\mathcal{B}_p$ e $\mathcal{B}_m$ de $\mathbb{R}^n$, $\mathbb{R}^p$ e $\mathbb{R}^m$. Defina $(f\circ g)(\boldsymbol{x}) = f\big(g(\boldsymbol{x})\big)$. Então a matriz correspondente a $f\circ g$ com respeito às bases $\mathcal{B}_n$ e $\mathcal{B}_m$ é

$$\boldsymbol{C} = \boldsymbol{A}\cdot\boldsymbol{B},$$

em que $c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}$.

Vamos ver agora algumas aplicações de transformações lineares.

2 Exercícios

Exercício 1 (Recordação). Considere o vetor do $\mathbb{R}^2$

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix},$$

escrito com relação à base (coordenadas canônicas)

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Quais são as coordenadas de $\boldsymbol{x}$ na base formada pelos vetores

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}?$$

E com relação à base formada pelos vetores

$$\boldsymbol{w}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \boldsymbol{w}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}?$$