MATRIZES SIMÉTRICAS TEM AUTOVALOR REAL

DANIEL V. TAUSK

Este é um texto complementar para o curso de MAT3458–Álgebra Linear II ministrado na Escola Politécnica da Universidade de São Paulo. No curso mostramos que todo operador linear simétrico num espaço vetorial real de dimensão finita munido de um produto interno pode ser diagonalizado em uma base ortonormal de autovetores. A demonstração desse resultado usa o seguinte teorema.

Teorema. *Se n é um inteiro positivo e $A \in M_n(\mathbb{R})$ é uma matriz real simétrica, então o polinômio característico de A possui uma raiz real.*

Existem duas abordagens para demonstrar esse Teorema: a primeira é mostrar que toda raiz complexa do polinômio característico de uma matriz real simétrica é real (não se esqueça que quando dizemos que um número é "complexo", não estamos excluindo a possibilidade de ele ser real). Aí, apelando para o Teorema Fundamental da Álgebra que diz que todo polinômio complexo não constante possui uma raiz complexa, obtemos a conclusão desejada. A segunda abordagem usa resultados de Cálculo Diferencial sobre máximos e mínimos de funções diferenciáveis. Nós discutiremos a segunda abordagem num apêndice e tratamos agora da primeira. A primeira abordagem é apresentada na maioria dos textos de Álgebra Linear usando a noção de produto interno em espaços vetoriais complexos. Apresentamos aqui uma demonstração que evita essa noção.

Se A é uma matriz complexa, denotamos por $\overline{A}$ o *complexo conjugado* de A, isto é, a matriz obtida de A pela tomada do complexo conjugado de todas as suas entradas. Segue facilmente do fato que as identidades

$$\overline{z+w} = \overline{z} + \overline{w} \quad \text{e} \quad \overline{zw} = \overline{z}\,\overline{w}$$

valem para quaisquer números complexos z e w que as identidades correspondentes

$$\overline{A+B} = \overline{A} + \overline{B} \quad \text{e} \quad \overline{AB} = \overline{A}\,\overline{B} \tag{1}$$

também valem para matrizes A e B, nos casos em que a soma $A+B$ e o produto AB fizerem sentido, respectivamente. Recordamos também que a transposta do produto de duas matrizes é igual ao produto das transpostas na ordem oposta, isto é, vale a identidade

$$(AB)^{\mathrm{t}} = B^{\mathrm{t}} A^{\mathrm{t}} \tag{2}$$

Date: 15 de outubro de 2019.

para matrizes quaisquer A e B tais que o produto AB faça sentido.

Estamos prontos agora para demonstrar o Teorema.

Seja λ uma raiz complexa do polinômio característico de A. Isso significa que a matriz $A - \lambda\mathrm{I}$ possui determinante nulo, em que I denota a matriz identidade $n \times n$. Segue daí que o sistema linear homogêneo

$$(A - \lambda\mathrm{I})v = 0 \tag{3}$$

possui uma solução não trivial, isto é, existe uma matriz complexa coluna $v \in M_{n\times 1}(\mathbb{C})$ não nula tal que a igualdade (3) é satisfeita (note que como a matriz de coeficientes $A - \lambda\mathrm{I}$ é complexa, garantimos apenas a existência de uma solução não trivial complexa, isto é, não podemos garantir *a priori* que v seja real). De (3) vem

$$Av = \lambda v$$

que é simplesmente a igualdade que normalmente é usada para definir autovetores e autovalores. Agora multiplicamos os dois lados dessa igualdade pela esquerda pela matriz linha que é a transposta do complexo conjugado de v obtendo

$$\overline{v^{\mathrm{t}}}Av = \lambda\overline{v^{\mathrm{t}}}v = \lambda\big(|v_1|^2 + \cdots + |v_n|^2\big) \tag{4}$$

em que $v_1, \ldots, v_n$ denotam as entradas da matriz v e $|z|$ denota o módulo de um número complexo z. Tendo em mente a identidade (2), tomamos agora a transposição dos dois lados da igualdade (4), o que nos dá:

$$v^{\mathrm{t}}A^{\mathrm{t}}\overline{v} = \lambda\big(|v_1|^2 + \cdots + |v_n|^2\big); \tag{5}$$

note que do lado direito de (4) temos uma matriz 1×1 que não foi afetada pela transposição. Tendo em mente a segunda identidade em (1), tomamos agora o complexo conjugado dos dois lados de (5) obtendo

$$\overline{v^{\mathrm{t}}}\;\overline{A^{\mathrm{t}}}\;v = \overline{\lambda}\big(|v_1|^2 + \cdots + |v_n|^2\big); \tag{6}$$

note que $|v_1|^2 + \cdots + |v_n|^2$ é real e portanto não é afetado pela conjugação complexa. Como A é real e simétrica, temos que $\overline{A^{\mathrm{t}}} = A$ e portanto comparando (4) e (6) vem:

$$\lambda\big(|v_1|^2 + \cdots + |v_n|^2\big) = \overline{\lambda}\big(|v_1|^2 + \cdots + |v_n|^2\big).$$

Como v é não nulo, podemos cancelar $|v_1|^2 + \cdots + |v_n|^2 > 0$ e daí concluímos que $\lambda = \overline{\lambda}$, isto é, que λ é real. Isso completa a demonstração do Teorema.

Apêndice: uma demonstração usando Cálculo

Seja $A \in M_n(\mathbb{R})$ uma matriz real e considere a função $f : \mathbb{R}^n \to \mathbb{R}$ definida por

$$f(x) = \langle Ax, x\rangle = x^{\mathrm{t}}Ax,$$

para todo $x \in \mathbb{R}^n$, em que denotamos também por x a matriz coluna com as mesmas coordenadas que x e por $\langle\cdot,\cdot\rangle$ o produto interno usual de $\mathbb{R}^n$. Uma fórmula mais explícita para f pode ser escrita usando somatórios, como segue:

$$f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j, \tag{7}$$

em que $A = (a_{ij})_{n\times n}$. Considere a esfera S em $\mathbb{R}^n$ de centro na origem e raio unitário, isto é:

$$S = \big\{x \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\big\} = \big\{x \in \mathbb{R}^n : \langle x, x\rangle = 1\big\}.$$

Recorde que o *Teorema de Weierstrass* nos diz que uma função contínua definida num conjunto compacto não vazio possui um ponto de máximo e um ponto de mínimo global. Como S é compacta e f é contínua, temos então que a restrição de f a S possui um ponto de máximo global, isto é, existe um ponto $x \in S$ tal que $f(x) \geq f(y)$, para todo $y \in S$ (pelo mesmo motivo, a restrição de f a S possui um ponto de mínimo global, mas o ponto de máximo é suficiente para nossos propósitos). Nós vamos mostrar agora que, se A for simétrica, então esse ponto x é um autovetor de A; seguirá daí que A possui um autovalor real, isto é, que o polinômio característico de A possui uma raiz real.

Usando a fórmula (7), podemos calcular as derivadas parciais de f como segue:

$$\begin{aligned}\frac{\partial f}{\partial x_k}(x) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\frac{\partial x_i}{\partial x_k}x_j + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i\frac{\partial x_j}{\partial x_k}\\ &= \sum_{j=1}^{n} a_{kj}x_j + \sum_{i=1}^{n} a_{ik}x_i;\end{aligned}$$

usamos aqui que $\frac{\partial x_i}{\partial x_k}$ é igual a 1 para $i = k$ e é igual a zero para $i \neq k$ (e o fato análogo para $\frac{\partial x_j}{\partial x_k}$). Segue daí que o gradiente da função f no ponto x é dado por:

$$\nabla f(x) = \Big(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\Big) = Ax + A^{\mathrm{t}}x.$$

No caso em que A é simétrica, temos:

$$\nabla f(x) = 2Ax. \tag{8}$$

Note que o ponto de máximo x cuja existência foi garantida pelo Teorema de Weierstrass não é um ponto de máximo local de $f : \mathbb{R}^n \to \mathbb{R}$, de modo que não podemos concluir que o gradiente de f seja nulo em x; temos apenas que x é um ponto de máximo da restrição de f a S, ou seja, um ponto de máximo de f sujeito ao vínculo

$$g(x) = 1,$$

em que $g : \mathbb{R}^n \to \mathbb{R}$ é definida por:

$$g(x) = x_1^2 + \cdots + x_n^2 = \langle x, x\rangle.$$

O gradiente de g pode ser calculado facilmente e o resultado é:

$$\nabla g(x) = 2x. \tag{9}$$

Como $\nabla g(x) \neq 0$, o *método dos multiplicadores de Lagrange* nos diz agora que existe $\lambda \in \mathbb{R}$ tal que

$$\nabla f(x) = \lambda \nabla g(x).$$

Tendo em vista (8) e (9), segue agora que $Ax = \lambda x$, ou seja, x é um autovetor de A com autovalor λ. Isso completa a demonstração do fato que o polinômio característico de uma matriz real simétrica possui uma raiz real.

Se você não tem familiaridade com o método dos multiplicadores de Lagrange, pode usar a seguinte abordagem alternativa: considere a função h definida por

$$h(x) = \frac{\langle Ax, x\rangle}{\langle x, x\rangle} = \frac{f(x)}{g(x)},$$

para todo $x \in \mathbb{R}^n$ com $x \neq 0$. Temos que h coincide com f na esfera S e que h é constante em qualquer semireta que sai da origem. Daí, se $x \in S$ é um ponto de máximo global da restrição de f a S, então x é um ponto de máximo global da função h (sem vínculo). Segue daí que:

$$\nabla h(x) = 0.$$

Calculamos agora o gradiente de h como segue:

$$\nabla h(x) = \frac{\nabla f(x) g(x) - f(x) \nabla g(x)}{g(x)^2};$$

usando (8), (9) e o fato que $g(x) = 1$, obtemos então:

$$\nabla h(x) = 2\big(Ax - \langle Ax, x\rangle x\big).$$

Como $\nabla h(x) = 0$, conclui-se que $Ax = \lambda x$, com $\lambda = \langle Ax, x\rangle$. Novamente, isso demonstra que x é um autovetor de A e que $\lambda = \langle Ax, x\rangle$ é um autovalor (real) de A.