1 Elementos de Álgebra Linear

O objetivo desse documento é fornecer elementos de álgebra linear para a compreensão de algoritmos relacionados à inteligência artificial que serão vistos no decorrer do curso de Introdução à Inteligência Artificial. Iniciamos recordando brevemente as operações básicas com matrizes de números reais. Em seguida, trataremos sobre autovalores e autovetores. Por fim, veremos o que é a decomposição em valores singulares de uma matriz e proporemos alguns exercícios em Python para fixar os conteúdos aprendidos.

1.1 Introdução

Dentro da área de inteligência artificial é necessário trabalhar com grandes quantidades de dados. Estes dados são representados geralmente por números reais e são armazenados em matrizes. Por esta razão é necessário o domínio das operações básicas envolvendo estes objetos matemáticos.

Uma matriz é uma tabela de números organizados em linhas e colunas. O exemplo a seguir mostra uma matriz com 3 linhas e duas colunas.

- **Exemplo 1:** Qual é a posição do número 8 na matriz abaixo?

$$A = \begin{pmatrix} -2 & 0 \\ 4 & -1 \\ 5 & 8 \end{pmatrix}.$$

◇ **Solução:** Podemos observar que o número 8 está na terceira linha e na segunda coluna da matriz A e pode ser denotado por a_{32}.

Matematicamente, uma matriz genérica A com m linhas e n colunas é dita ter tamanho $m \times n$ e seus elementos são representados por a_{ij}, onde os índices i e j indicam, respectivamente, a linha e a coluna da matriz em que o elemento se encontra. Podemos escrever então que $A = (a_{ij})_{m\times n}$. Se $m = n$, dizemos que a matriz é quadrada ou de ordem n. Vamos assumir neste documento que as entradas de uma matriz são constituídas por números reais.

1.2 Operações com matrizes

As principais operações realizadas com matrizes são:

1. Multiplicação por escalar: dado $\alpha \in \mathbb{R}$ e $A = (a_{ij})_{m\times n}$, temos que $\alpha A = (\alpha a_{ij})_{m\times n}$, isto é, cada elemento de A é multiplicado por α na nova matriz.

2. Adição: dadas $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{m\times n}$, temos que
$$A + B = (a_{ij} + b_{ij})_{m\times n}.$$
Isto significa que somente podemos somar ou subtrair matrizes com o mesmo número de linhas e colunas.

3. Multiplicação de matrizes: dadas $A = (a_{ij})_{m\times n}$ e $B = (b_{ij})_{n\times p}$, o produto $C = AB$ é uma nova matriz de tamanho $m \times p$ cujos elementos são dados por
$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$
Note que esta operação só é possível se o número de colunas da matriz A for igual ao número de linhas da matriz B.

- **Exemplo 2:** Dadas as matrizes A, B e C abaixo, verifique se são possíveis as operações AB, BA, $3C$, $A+B$ e BC. Em caso afirmativo, obtenha o resultado de cada operação.

◇ **Solução:**

$$A = \begin{pmatrix} -2 & 0 \\ 4 & -1 \\ 5 & 8 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & -2 \\ 3 & 2 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 1 \\ 0 & 5 \end{pmatrix}.$$

a) É possível efetuar o produto AB pois A é de tamanho 3×2 e B é de tamanho 2×3:
$$AB = \begin{pmatrix} -2 & 0 & 4 \\ 1 & -2 & -14 \\ 29 & 16 & 38 \end{pmatrix}.$$

b) É possível efetuar o produto BA pois B é de tamanho 2×3 e A é de tamanho 3×2:
$$BA = \begin{pmatrix} -12 & -16 \\ 32 & 46 \end{pmatrix}.$$
Os itens a) e b) são interessantes pois nos mostram que o produto de duas matrizes não é comutativo, ou seja, em geral AB não é igual a BA.

c) É possível realizar a operação $3C$. Basta multiplicar cada elemento de C por 3:
$$3C = \begin{pmatrix} 6 & 3 \\ 0 & 15 \end{pmatrix}.$$

d) Não é possível realizar as operações $A + B$ e BC. No caso da adição, as matrizes envolvidas são de tamanhos diferentes. No caso da multiplicação, o número de colunas de B é diferente do número de linhas de C.

1.3 Alguns tipos de matrizes

Alguns tipos de matrizes aparecem com frequência em problemas envolvendo inteligência artificial e serão discutidos brevemente a seguir.

1. Matriz transposta: dada $A = (a_{ij})_{m\times n}$, a transposta de A, denotada por A^T, é dada por $A^T = (a_{ji})_{n\times m}$, ou seja, cada linha na matriz original se torna coluna na matriz transposta. Considerando a matriz A do Exemplo 1:

$$\text{se} \quad A = \begin{pmatrix} -2 & 0 \\ 4 & -1 \\ 5 & 8 \end{pmatrix}, \quad \text{então} \quad A^T = \begin{pmatrix} -2 & 4 & 5 \\ 0 & -1 & 8 \end{pmatrix}.$$

2. Matriz diagonal: uma matriz A quadrada é chamada diagonal se $a_{ii} \neq 0$ e $a_{ij} = 0$ para todos i, j. Isto é, apenas os elementos que estão na diagonal de A são diferentes de zero. Por esta razão é comum escrever uma matriz diagonal de ordem n pela notação compacta $A = \text{diag}(a_{11}, a_{22}, \ldots, a_{nn})$. Considere, como exemplo, as matrizes:

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{e} \quad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ambas são matrizes diagonais de ordem 3, a matriz I tem o nome especial de matriz identidade, pois possui apenas o número 1 em sua diagonal principal.

3. Matriz simétrica: uma matriz A quadrada é chamada de simétrica se $a_{ij} = a_{ji}$, ou seja, $A^T = A$. Considere a matriz

$$A = \begin{pmatrix} 6 & 3 \\ 3 & 6 \end{pmatrix}.$$

Note que a transposta de A é igual a ela própria. Portanto, A é simétrica.

1.4 Determinantes e matrizes inversas

Um determinante é um número real associado a uma matriz quadrada denotado por $\det(A)$ ou $|A|$. Para calcular o determinante de qualquer matriz $n \times n$, é preciso aplicar o desenvolvimento em cofatores, partindo de qualquer linha (ou coluna) da matriz. No caso de matrizes de ordem 2 ou 3, este desenvolvimento resulta em fórmulas bastante conhecidas:

a) $n = 2$: se $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, então $\det(A) = a_{11}a_{22} - a_{12}a_{21}$.

b) $n = 3$: se $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, então

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

A inversa de uma matriz quadrada A de ordem n é definida como sendo a única matriz X, também de ordem n, que satisfaz a equação matricial $AX = XA = I$, em que I é a matriz identidade $n \times n$. Costumamos denotar a inversa de uma matriz A por A^{-1}. É possível demonstrar que uma matriz admite inversa se, e somente se, seu determinante for diferente de zero. Assim, se A é uma matriz invertível, usando propriedades dos determinantes, conseguimos mostrar que

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

- **Exemplo 3:** Calcule o determinante da matriz $A = \begin{pmatrix} 3 & -2 \\ 1 & 4 \end{pmatrix}$.

◇ **Solução:** O valor de seu determinante é $\det(A) = 3 \cdot 4 - (-2) \cdot 1 = 12 + 2 = 14$. Como $\det(A) \neq 0$, A é uma matriz invertível e o determinante da inversa é igual a $\det(A^{-1}) = 1/14$.

1.5 Autovalores e autovetores

Seja A uma matriz de ordem n. Se existirem um escalar $\lambda \in \mathbb{R}$ e um vetor $v \in \mathbb{R}^n$ não nulo tais que $Av = \lambda v$, diremos que λ é um autovalor e v é um autovetor de A. Esta definição nos permite escrever:

$$Av - \lambda v = 0 \Leftrightarrow (A - \lambda I)v = 0.$$

Note que a última equação é um sistema linear homogêneo. Como, pela definição, um autovetor não pode ser nulo, o sistema terá necessariamente ininitas soluções. Portanto, recordando a teoria de sistemas de equações lineares, para que um sistema homogêneo tenha infinitas soluções, o determinante da matriz de coeficientes do sistema precisa ser igual a zero. Logo,

$$\det(A - \lambda I) = 0.$$

Esta última equação é chamada de polinômio característico da matriz A, pois o resultado do determinante será um polinômio de grau n na varivável λ, denotado por $p(\lambda)$.

- **Exemplo 4:** Calcule os autovalores e autovetores da matriz $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.

◇ **Solução:** Montando a matriz $A - \lambda I$ obtemos

$$A - \lambda I = \begin{pmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{pmatrix}.$$

O polinômio característico de A é dado pelo determinante da matriz anterior, ou seja,

$$p(\lambda) = \det(A - \lambda I) = (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 3.$$

Os autovalores de A são dados pelas raízes da equação $p(\lambda) = 0$:

$$p(\lambda) = 0 \Rightarrow \lambda^2 - 4\lambda + 3 = 0 \Rightarrow \lambda = 1 \quad \text{ou} \quad \lambda = 3.$$

Os autovetores associados a cada autovalor são obtidos resolvendo o sistema linear homogêneo $(A - \lambda I)v = 0$ para cada valor de λ encontrado. Para o caso $\lambda = 1$ temos:

$$(A - I)v = 0 \Rightarrow \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \left\{ \begin{array}{rcl} x_1 + x_2 & = & 0 \\ x_1 + x_2 & = & 0 \end{array} \right. .$$

Isolando x_2 em qualquer uma das equações do sistema obtemos $x_2 = -x_1$. Fazendo $x_1 = 1$, tem-se que $x_2 = -1$. Assim, o autovetor procurado é $v_1 = (1, -1)^T$. Para o caso $\lambda = 3$ temos:

$$(A - 3I)v = 0 \Rightarrow \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \left\{ \begin{array}{rcl} -x_1 + x_2 & = & 0 \\ x_1 - x_2 & = & 0 \end{array} \right. .$$

Isolando x_2 em qualquer uma das equações obtemos $x_2 = x_1$. Fazendo $x_1 = 1$, segue que $x_2 = 1$. Logo, o autovetor procurado é $v_2 = (1, 1)^T$.

Algumas observações importantes sobre cálculo de autovalores e autovetores são:

a) Se a matriz A é grande, resolver a equação $p(\lambda) = 0$ para encontrar seus autovalores pode não ser uma tarefa fácil. Por esta razão, métodos numéricos que evitam a resolução de $p(\lambda) = 0$ são mais indicados.

b) Na maioria das aplicações em que é necessário encontrar um autovetor, estamos mais interessados na direção que ele fornece ao invés de sua magnitude. Portanto, é comum deixá-los com tamanho unitário durante os cálculos. Chamamos este procedimento de normalização. Para normalizar os autovetores encontrados no Exemplo 4, basta multiplicar cada uma de suas componentes pelo inverso de sua norma euclidiana. Lembrando que a norma euclidiana de um vetor é dada por

$$\|v\|_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2},$$

é fácil ver que $\|v_1\|_2 = \|v_2\|_2 = \sqrt{2}$. Portanto, os autovetores normalizados são

$$w_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \quad \text{e} \quad w_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

Retornando à matriz A do Exemplo 4, seja V a matriz formada pelos autovetores normalizados de A dispostos em colunas:

$$V = (w_1 \mid w_2) = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Note que $VV^T = V^TV = I$. Ou seja, a inversa de V é igual a sua transposta. V então é chamada de matriz ortogonal. Fazendo $\Lambda = \text{diag}(\lambda_1, \lambda_2) = \text{diag}(1, 3)$, podemos escrever:

$$AV = V\Lambda \Rightarrow A = V\Lambda V^T = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Esta forma de decompor a matriz A é chamada de decomposição em autovalores e autovetores (ou decomposição espectral). É importante observar que nem toda matriz quadrada pode ser decomposta dessa forma. No entanto, toda matriz simétrica (como é o caso da matriz do Exemplo 4) possui uma decomposição deste tipo. Além disso, é possível provar que os autovalores de uma matriz simétrica são sempre números reais.

1.6 Decomposição em valores singulares

Qualquer matriz A de tamanho $m \times n$ pode ser decomposta como um produto de três matrizes da seguinte forma

$$A = U \Sigma V^T.$$

Esta decomposição, denominada de decomposição em valores singulares, satisfaz as condições:

1. U é uma matriz de tamanho $m \times n$ cujas colunas são constituídas por vetores dois a dois ortogonais. Estes vetores, denotados por u_i, são chamados de vetores singulares à esquerda de A.

2. Σ é uma matriz diagonal $n \times n$ que contém os números reais $\sigma_1 \geq \sigma_2 \geq \dots \sigma_n \geq 0$ chamados de valores singulares de A, isto é, $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$.

3. V é uma matriz ortogonal de ordem n cujas colunas são constituídas por vetores v_i chamados de vetores singulares à direita de A.

A decomposição em valores singulares nos permite escrever a matriz A da forma

$$A = \sum_{i=1}^{n} \sigma_i u_i v_i^T,$$

onde cada termo $u_i v_i^T$ é uma matriz quadrada de ordem n, pois u_i tem tamanho $n \times 1$ e v_i^T tem tamanho $1 \times n$. Para obter a decomposição em valores singulares de A é necessário utilizar os seguintes resultados:

1. Os autovalores da matriz $A^T A$ são σ_i^2 e os autovetores correspondentes são os vetores v_i (vetores singulares à direita).

2. Os autovalores da matriz AA^T são σ_i^2 e os autovetores correspondentes são os vetores u_i (vetores singulares à esquerda).

Os dois resultados anteriores nos mostram que o cálculo de valores/vetores singulares de A está diretamente relacionado ao cálculo de autovalores/autovetores das matrizes $A^T A$ e AA^T.

As figuras a seguir ilustram a interpretação geométrica da decomposição em valores singulares de uma matriz. Para construí-las, utilizamos a matriz A de ordem 2 vista no Exemplo 4. A Figura 1a) mostra a circunferência unitária centrada em $(0, 0)$ com destaque para dois vetores ortogonais, também de tamanho unitário. Um dos vetores da figura foi denotado por w. A Figura 1b) mostra o efeito da aplicação da matriz V^T sobre todos os pontos da circunferência. Observe que este efeito fez com que os vetores destacados fossem rotacionados no sentido horário. Assim, o vetor w foi transformado no vetor $V^T w$. A Figura 1c) mostra o efeito da aplicação da matriz Σ sobre os vetores rotacionados. Note que os valores singulares fazem com que os vetores tenham o seu tamanho alterado. Neste caso, o vetor $V^T w$ foi transformado no vetor $\Sigma V^T w$. Por fim, a Figura 1d) mostra o efeito de aplicar a matriz U sobre os vetores transformados da figura c). Neste caso, o vetor $\Sigma V^T w$ foi transformado no vetor $U \Sigma V^T w = Aw$. Podemos notar por esta figura que a circunferência unitária foi transformada em uma elipse.

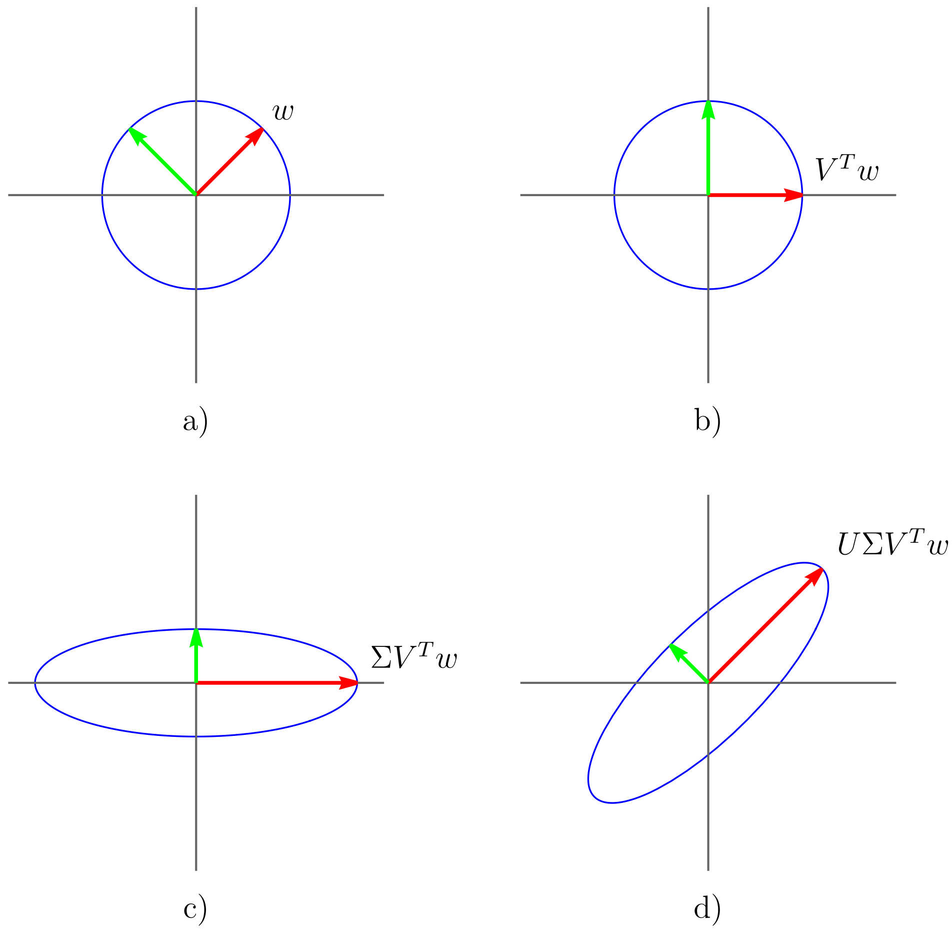

Figura 1: Interpretação geométrica da decomposição em valores singulares.

Encerramos o estudo desses tópicos de Álgebra Linear com uma aplicação muito comum da decomposição em valores singulares: a compressão de imagens. Uma imagem pode ser representada por uma matriz A cujos elementos são números reais. Uma vez que a decomposição em valores singulares de A é obtida, podemos utilizar os primeiros k valores e vetores singulares dessa decomposição para obter uma nova matriz $\hat{A}$ que representa uma aproximação para a imagem original. Assim, a matriz $\hat{A}$ é dada por

$$\hat{A} = \sum_{i=1}^{k} \sigma_i u_i v_i^T.$$

A vantagem desse processo é evitar o armazenamento de grandes quantidades de dados que representam a imagem original, sendo que uma imagem também nítida pode ser obtida com um armazenamento de dados muito menor. Como exemplo para ilustrar esse procedimento, considere a Figura 2 abaixo. A Figura 2 a) mostra uma imagem em escala cinza da cidade de Nova Iorque. A matriz numérica que representa essa imagem possui 398 linhas e 642 colunas, resultando em um total de $398 \times 642 = 255516$ números para serem armazenados. Já a Figura 2 b) ilustra uma aproximação para a imagem original em que foram utilizados os 30 primeiros valores e vetores singulares da decomposição de A. Neste caso, a quantidade de dados a serem armazenados são:

- ◇ $398 \times 30 = 11940$ números relacionados aos vetores u_i,
- ◇ 30 valores singulares,
- ◇ $30 \times 642 = 19260$ números relacionados aos vetores v_i,

totalizando 31230 números (aproximadamente 12% do total de dados).

a) Imagem original.

b) Imagem reconstruída.

Figura 2: Compressão de imagens via decomposição em valores singulares.

Bibliografia

1. Applied Numerical Linear Algebra - J. W. Demmel - SIAM, 1997.
2. Matrix Analysis and Applied Linear Algebra - C. D. Meyer - SIAM, 2000.
3. Matrizes, Vetores e Geometria Analítica - R. J. Santos - UFMG, 2007.