www.datascienceacademy.com.br

Matemática Para Machine Learning

Vetores, Matrizes e Algebra Linear

Os matemáticos definem **vetor** como um membro de um espaço vetorial, mas utilizaremos uma definição mais concreta: um vetor é uma sequência ordenada de valores. Por exemplo, em um espaço bidimensional, temos vetores como **x** = 〈3, 4〉 e **y** = 〈0, 2〉 . Seguimos a convenção habitual de usar caracteres em negrito para representar nomes de vetores, embora alguns autores utilizem setas ou barras sobre os nomes **x** ou **y**.

Os elementos de um vetor podem ser acessados com a utilização de subscritos:

$$\mathbf{z} = \langle z_1, z_2, \ldots, z_n \rangle$$

Um ponto confuso: IA é um trabalho sintético de muitos subcampos, que de diferentes maneiras chamam seus vetores de sequências, listas ou tuplas e, frequentemente, utilizam as notações 〈1, 2〉 , [1, 2], ou (1, 2).

As duas operações fundamentais sobre vetores são a adição vetorial e a multiplicação escalar. A adição vetorial x + y é a soma dos elementos correspondentes dos vetores:

$$x + y = \langle 3 + 0, 4 + 2 \rangle = \langle 3, 6 \rangle$$

A multiplicação escalar multiplica cada elemento por uma constante:

$$5x = \langle 5 \times 3, 5 \times 4 \rangle = \langle 15, 20 \rangle$$

O comprimento de um vetor é indicado por |x| e é calculado tomando-se a raiz quadrada da soma dos quadrados dos elementos: $|\mathrm{x}| = \sqrt{(3^2 + 4^2)} = 5$. O produto de ponto (também chamado produto escalar ou dot product) de dois vetores x · y é a soma dos produtos dos elementos correspondentes, isto é, $x \cdot y = \Sigma_i \, x_i \, y_i$ ou, em nosso caso específico:

$$x \cdot y = 3 \times 0 + 4 \times 2 = 8$$

Os vetores frequentemente são interpretados como segmentos de reta orientados (setas) em um espaço euclidiano n dimensional. Então, a adição vetorial é equivalente a conectar o final de um vetor ao início do outro e o produto escalar x · y é igual a |x| |y| cos θ, onde θ é o ângulo entre x e y.

Uma **matriz** é um array retangular de valores organizados em linhas e colunas. Aqui temos uma matriz A de tamanho 3 × 4:

$$\begin{pmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \mathbf{A}_{1,3} & \mathbf{A}_{1,4} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \mathbf{A}_{2,3} & \mathbf{A}_{2,4} \\ \mathbf{A}_{3,1} & \mathbf{A}_{3,2} & \mathbf{A}_{3,3} & \mathbf{A}_{3,4} \end{pmatrix}$$

O primeiro índice de Ai,j especifica a linha e o segundo especifica a coluna. Em linguagem de programação, Ai,j frequentemente é escrito como A[i,j] ou A[i][j].

A soma de duas matrizes é definida pela adição de elementos correspondentes; desse modo, (A+B)i,j = Ai,j + Bi,j (a soma é indefinida se A e B têm tamanhos diferentes). Também podemos definir a multiplicação de uma matriz por um escalar: (cA)i,j = cAi,j. A multiplicação de matrizes (o produto de duas matrizes) é mais complicada. O produto AB é definido apenas se A tem o tamanho a × b e B tem o tamanho b × c (**isto é, a segunda matriz tem um número de linhas igual ao número de colunas da primeira matriz**); o resultado é uma matriz de tamanho a × c. Se as matrizes tiverem tamanho apropriado, o resultado será:

$$(\mathbf{AB})_{i,k} = \sum_{j} \mathbf{A}_{i,j}\mathbf{B}_{j,k}$$

A multiplicação de matrizes não é comutativa, mesmo para matrizes quadradas: AB ≠ BA em geral. No entanto, é associativa: (AB)C = A(BC). Observe que o produto escalar pode ser expresso em termos de uma transposição e uma multiplicação de matrizes: x · y = xTy.

A matriz identidade I tem elementos Ii,j iguais a 1 quando i = j e iguais a 0 em caso contrário. Ela tem a propriedade de que AI = A para todo A. A transposta de A, escrita como AT é formada transformando-se as linhas em colunas e vice-versa ou, de modo mais formal, por ATi, j = Aj,i. O inverso de uma matriz quadrada A é outra matriz quadrada A−1 tal que A−1 A = I.

As matrizes são usadas para resolver sistemas de equações lineares. Considere o conjunto de equações a seguir, para o qual queremos encontrar uma solução em x, y e z:

+2x + y – z = 8
–3x – y + 2z = −11
–2x + y + 2x = –3

Podemos representar esse sistema como a equação matricial A x = b, onde

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & -1 \\ -3 & -1 & 2 \\ -2 & 1 & 2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 8 \\ -11 \\ -3 \end{pmatrix}$$

Para resolver A x = b, multiplicamos ambos os lados por A−1, produzindo A−1Ax = A−1b, que, simplificando, dá x = A−1b. Depois de inverter A e multiplicar por b, obtemos a resposta:

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$$

Referência:

Artificial Intelligence: A Modern Approach
https://www.amazon.com.br/Artificial-Intelligence-Modern-Approach-Global-ebook/dp/B01HEY2P66/ref=tmm_kin_swatch_0?_encoding=UTF8&qid=1536709293&sr=8-2-fkmr0