Data Science Academy

www.datascienceacademy.com.br

Matemática Para Machine Learning

Sistemas de Equações Lineares e Machine Learning

Um sistema de equações lineares (ou sistema linear) é uma coleção de duas ou mais equações lineares envolvendo o mesmo conjunto de variáveis. Vamos considerar esse sistema linear de 3 equações e 3 incógnitas:

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}$$

Esse sistema pode ser representado por sua forma matricial, assim:

$$\overbrace{\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}}^{A} \overbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}^{X} = \overbrace{\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}}^{D}$$

Ou apenas: Ax = D

A matriz A é chamada de matriz de coeficientes, a matriz colunar x (ou vetor) é chamada de matriz de variáveis e a matriz colunar D é chamada de matriz dos termos independentes. Podemos unir as Matrizes A e D criando uma matriz aumentada (que estudamos nos capítulos anteriores), que pode ser representada da seguinte forma:

$$[A|D] = \begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{bmatrix}$$

Podemos também representar a matriz acima em sua forma vetorial, da seguinte forma:

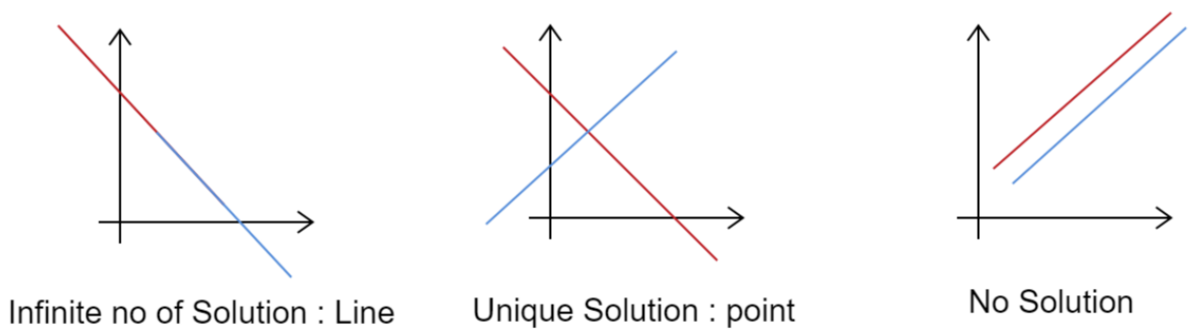
$$x\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + y\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + z\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

A forma como você representa um sistema linear, depende do tipo de operação que você deseja fazer com o sistema.

Cada sistema linear pode ter apenas um dos três possíveis conjuntos de soluções:

- O sistema tem uma solução única e exclusiva.
- O sistema tem muitas infinitas soluções.
- O sistema não tem solução.

Para um sistema de duas variáveis (x e y), cada equação linear determina uma linha no plano xy. O conjunto de soluções é a interseção dessas linhas e, portanto, é uma linha, um único ponto ou não tem nenhum ponto em comum.



Hiperplano é um termo muito importante no aprendizado de máquina e é usado com muita frequência. Abaixo está uma boa explicação do hiperplano:

Um hiperplano é um subespaço cuja dimensão é uma a menos que a do seu espaço de ambiente. Se um espaço é tridimensional, então seus hiperplanos são planos bidimensionais, enquanto se o espaço é bidimensional, seus hiperplanos são as linhas unidimensionais. Esta noção pode ser usada em qualquer espaço geral em que o conceito da dimensão de um subespaço é definido.

Quando trabalhamos com algoritmos como o SVM (Support Vector Machines) por exemplo, os dados são movimentados de um hiperplano para outro, de modo que o algoritmo possa encontrar uma solução que classifique os dados (por exemplo).

Cada equação linear determina um hiperplano no espaço n-dimensional, onde n é o número de variáveis. O conjunto de soluções é a interseção desses hiperplanos.

Outras Características de um Sistema Linear

Um sistema linear é considerado consistente se tiver pelo menos uma solução e é considerado inconsistente se não tiver solução.

Um sistema linear é considerado independente se nenhuma das equações puder ser escrita como uma combinação linear de outras. Por exemplo, as equações x + y = 2 e 2x + 2y = 4 não são linearmente independentes, pois a segunda equação pode ser obtida multiplicando-se 2 pela 1ª equação.

A extensão de um conjunto de vetores é o conjunto de pontos obtidos por todas as combinações lineares do conjunto de vetores. Por exemplo, o conjunto de vetores {(1,0,0), (0,1,0), (0,0,1)} abrange todo o espaço de coordenadas real de 3 dimensões (R^3). Qualquer número n de vetores linearmente independentes com números reais pode abranger um espaço de coordenadas real de dimensão n. O intervalo de um determinado conjunto de vetores pode ser determinado pelos vetores linearmente independentes no conjunto.

O número máximo de linhas linearmente independentes de uma matriz é chamado de classificação de linha (Row Rank), e o número máximo de colunas linearmente independentes é chamado de classificação de coluna (Column Rank) da matriz.

Para qualquer matriz A:

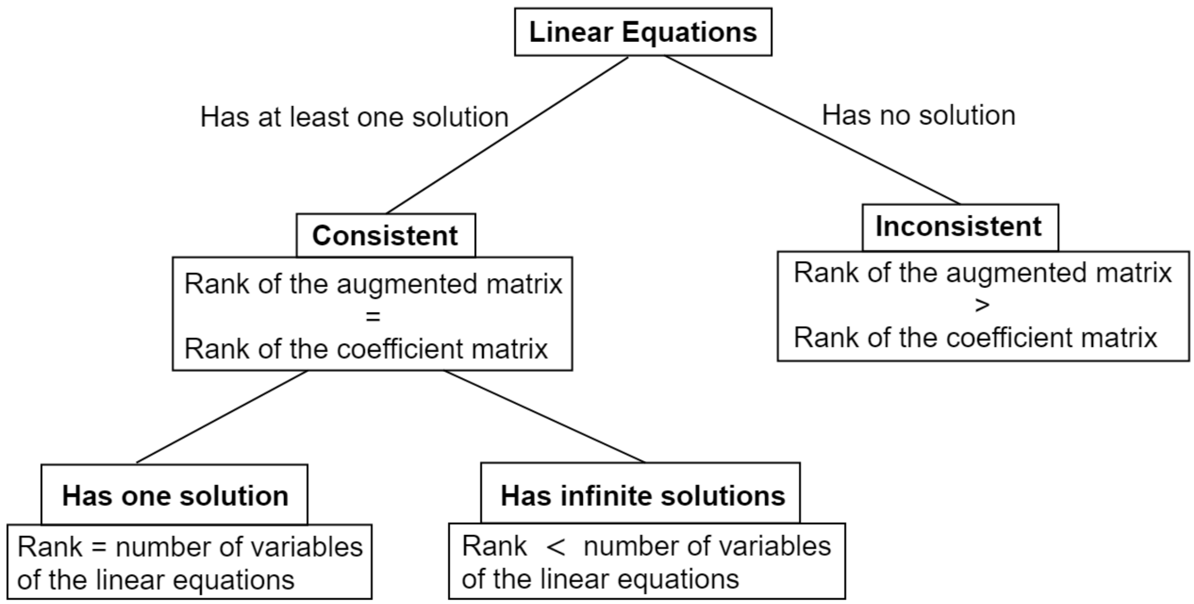
Row Rank de A = Column Rank de A = Rank de A

O Rank da matriz pode ser encontrado reduzindo-se a matriz em "Echelon Form's" (Forma de Echelon de Linha ou Forma de Echelon de Linha Reduzida). Isso é feito por operações de linha elementares e todo o método é chamado Eliminação de Gauss ou eliminação de Gauss-Jordan.

$$A_{ref} = \begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 0 & 1 & 4 & 3 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad A_{rref} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Row Echelon Form(ref) and Reduced Row Echelon Form (rref)

Se soubermos o rank de coeficientes e matrizes aumentadas, então podemos determinar o número de soluções de um sistema linear de equações. Abaixo está um diagrama que irá ajudá-lo a entender todo o conceito:



Existem muitas maneiras diferentes de resolver um sistema linear, como a Regra de Cramer, o Escalonamento e o método de Eliminação de Gauss e o método da matriz inversa.

Os sistemas lineares são usados principalmente com os autovalores e os autovetores, tema do próximo capítulo.

Referências:

What is the intuitive meaning of the basis of a vector space and the span?
https://math.stackexchange.com/questions/313504/what-is-the-intuitive-meaning-of-the-basis-of-a-vector-space-and-the-span

Rank of a Matrix
https://people.math.osu.edu/costin.9/264H/Rank.pdf

Echelon Form of a Matrix
https://stattrek.com/matrix-algebra/echelon-form.aspx

Eliminação Gaussiana
https://en.wikipedia.org/wiki/Gaussian_elimination