数学基础—线性代数

主要内容

- 线性代数（上）
  - 矩阵及其基本运算
    - 1.矩阵的基本概念与意义以及常见特殊矩阵
    - 2.矩阵的加法减法数乘以及性质
    - 3.矩阵的乘法以及性质
    - 4.矩阵运算在深度学习中的应用(初级)
    - 5.矩阵的迹，矩阵的转置，对称矩阵(协方差矩阵)
  - 矩阵的行列式
    - 1.行列式的引入
    - 2.行列式的计算
    - 3.特殊矩阵的行列式与行列式的性质
    - 4.行列式按行(列)展开，代数余子式
    - 5.行列式在接线性方程组中的应用：克莱姆法则
  - 矩阵的逆
    - 1.矩阵逆的引入以及矩阵逆的定义
    - 2.矩阵逆的计算
    - 3.矩阵逆的常用性质，以及特殊矩阵的逆
    - 4.矩阵逆在机器学习线性回归算法中的运用(初级)
    - 5.分块矩阵

矩阵的基本概念及意义以及常见的特殊矩阵

什么是矩阵

方阵

行向量

列向量

两个矩阵相等

零矩阵O

定义 由$m \times n$个数$a_{ij}(i = 1,2,\cdots,m; j = 1,2,\cdots,n)$排成的$m$行$n$列的数表

$$\begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{matrix}$$

称为m行n列矩阵，简称$m \times n$矩阵。为表示它是一个整数，总是加一个括弧，并用大写黑体字母表示它，记作

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

这$m \times n$个数称为矩阵A的元素，简称为元，数a_{ij}位于矩阵A的第i行第j列，称为矩阵A的(i,j)元. 以数a_{ij}为(i,j)元的矩阵可简记作(a_{ij})或$(a_{ij})_{m\times n}$。 m $\times$ n矩阵A也记作$A_{m\times n}$。

矩阵的基本概念及意义以及常见的特殊矩阵

一种线性变换

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ \cdots\cdots\cdots\cdots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n, \end{cases}$$

表示一个从变量 $x_1, x_2, \cdots, x_n$ ，到变量 $y_1, y_2, \cdots, y_m$ 的**线性变换**，其中 a_{ij} 为常数.

线性变换的系数 a_{ij} 构成矩阵 $A = (a_{ij})_{m\times n}$.

$$\begin{cases} y_1 = x_1, \\ y_2 = x_2, \\ \cdots\cdots\cdots\cdots \\ y_m = x_n \end{cases}$$

单位矩阵E

$$\begin{cases} y_1 = \lambda_1 x_1, \\ y_2 = \lambda_2 x_2, \\ \cdots\cdots\cdots\cdots \\ y_m = \lambda_n x_n \end{cases}$$

对角矩阵

矩阵的加法减法数乘以及性质

加法与数乘

定义 设有两个$m \times n$矩阵$A = (a_{ij})$和$B = (b_{ij})$, 那么矩阵A与B的和记作$A + B$，规定为

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

应该注意，只有当两个矩阵是同型矩阵时，这两个矩阵才能进行加法运算。

加法运算律

设A, B, C都是 $m * n$的矩阵

$$(A+B) = (B+A)$$

$$(A+B)+C = A+(B+C)$$

定义 数λ与矩阵A的乘积记作λA或$A\lambda$，规定为

$$\lambda A = A\lambda = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{bmatrix}$$

数乘运算律

数乘矩阵满足下列运算规律

（设A，B为$m \times n$矩阵，λ，μ为数）：

(i) $(\lambda\mu)A = \lambda(\mu A)$;

(ii) $(\lambda + \mu)A = \lambda A + \mu A$;

(iii) $\lambda(A+B) = \lambda A + \lambda B$.

矩阵的乘法以及性质

矩阵的乘法

定义 设 $A=(a_{ij})$是一个 $m\times s$矩阵，$B=(b_{ij})$是个 $s\times n$矩阵，那么规定矩阵A 与矩阵B的乘积是一个$m\times n$ 矩阵 $C=(c_{ij})$ ，其中

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{is}b_{sj}=\sum_{k=1}^{s}a_{ik}b_{jk}$$

$$(i=1,2,\cdots,m;\ j=1,2,\cdots,n),$$

并把次乘积记作 $C=AB$

$$(i)(AB)C=A(BC);$$

$$(ii)\lambda(AB)=(\lambda A)B=A(\lambda B);$$

$$(iii)A(B+C)=AB+AC,$$

$$(B+C)A=BC+CA.$$

$A=\begin{pmatrix}-2 & 4\\ 1 & -1\end{pmatrix}$ 与 $B=\begin{pmatrix}2 & 4\\ -3 & -6\end{pmatrix}$

AB 及 BA

$$AB=\begin{bmatrix}-2 & 4\\ 1 & -1\end{bmatrix}\begin{bmatrix}2 & 4\\ -3 & -6\end{bmatrix}=\begin{bmatrix}-16 & -32\\ 5 & 10\end{bmatrix}$$

$$BA=\begin{bmatrix}2 & 4\\ -3 & -6\end{bmatrix}\begin{bmatrix}-2 & 4\\ 1 & -1\end{bmatrix}=\begin{bmatrix}0 & 4\\ 0 & -6\end{bmatrix}$$

重点：

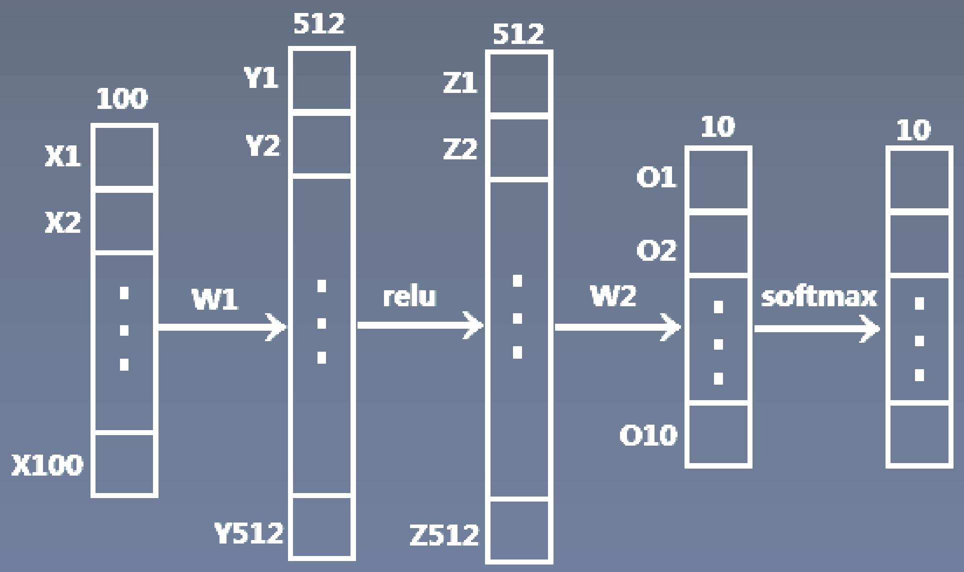
矩阵的乘法不满足交换律

矩阵运算在深度学习中的应用(初级)

数字图片识别

输入一张为数字(0-9)的图片，大小为10*10

下面也可以体现出矩阵是一种特征空间的变换



$W1$: $100 * 512$ 的矩阵
$W2$: $512 * 10$ 的矩阵

单样本：

$$(x_1, x_2, \cdots, x_{100})w_1 = (y_1, y_2, \cdots, y_{512})$$

$$(z_1, z_2, \cdots, z_{512})w_2 = (o_1, o_2, \cdots, o_{10})$$

n个样本：

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1,100} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{n100} \end{bmatrix} w_1 = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1,512} \\ \vdots & \vdots & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{n,512} \end{bmatrix}$$

$$\begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1,512} \\ \vdots & \vdots & & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{n512} \end{bmatrix} w_2 = \begin{bmatrix} o_{11} & o_{12} & \cdots & o_{1,10} \\ \vdots & \vdots & & \vdots \\ o_{n1} & o_{n2} & \cdots & o_{n,10} \end{bmatrix}$$

矩阵的迹，矩阵的转置，对称矩阵(协方差矩阵)

矩阵的迹

定义 在线性代数中，一个$n \times n$矩阵A的主对角线（从左上方至右下方的对角线）上各个元素的总和被称为矩阵A的迹（或迹数），一般记作$tr(A)$。

$$tr(A) = \sum_{i=1}^{n} a_{ii}$$

$tr(AB) = tr(BA)$对于满足矩阵乘法条件（型号匹配的）任意$A_{m\times n}$、$B_{n\times m}$均成立.

$(AB)_{m\times m}$ $(BA)_{n\times n}$

$$tr(AB) = \sum_{i=1}^{m} c_{ii} = \sum_{i=1}^{m}\sum_{s=1}^{n} a_{is} b_{si}$$

$$tr(BA) = \sum_{i=1}^{n} d_{ii} = \sum_{i=1}^{n}\sum_{s=1}^{m} b_{is} a_{si} = \sum_{s=1}^{m}\sum_{i=1}^{n} a_{si} b_{is} = \sum_{i=1}^{m}\sum_{s=1}^{n} a_{is} b_{si}$$

矩阵的迹，矩阵的转置，对称矩阵(协方差矩阵)

矩阵的转置

定义 把矩阵A的行换成同序数的列得到一个新矩阵，叫做矩阵A的**转置**，记作 A^T.

例如

$$A=\begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 1 \end{pmatrix}$$

的转置为

$$A^T=\begin{pmatrix} 1 & 3 \\ 2 & -1 \\ 0 & 1 \end{pmatrix}$$

运算律

$(i)(A^T)^T = A$

$(ii)(A+B)^T = A^T + B^T;$

$(iii)(\lambda A)^T = \lambda A^T$

$(iiii)(AB)^T = B^T A^T$

矩阵的迹，矩阵的转置，对称矩阵(协方差矩阵)

对称矩阵

定义 设A为n阶方阵，如果满足 $A^T=A$ ，即 $a_{ij}=a_{ji}(i,j=1,2,\cdots,n)$,
那么A称为**对称矩阵**，简称**对称阵**.

对称阵的特点是：它的元素以对角线为对称轴对应相等
很显然单位矩阵以及对角矩阵都为对称矩阵

例8 设矩阵 $X=(x_1,x_2,\cdots,x_n)^T$ 满足 $X^TX=1$，E为n阶单位阵，$H=E-2XX^T$
证明H是对称阵，且 $HH^T=E$.

$$HH^T=(E-2XX^T)(E-2XX^T)^T=(E-2XX^T)(E^T-(2XX^T)^T)$$

$$=(E-2XX^T)\ (E-2XX^T)=E-2XX^T-2XX^T+4XX^TXX^T$$

$$=E$$

矩阵的迹，矩阵的转置，对称矩阵(协方差矩阵)

协方差矩阵

N个样本，每个样本的特征的维度为n
易证协方差矩阵是对称矩阵(这个结论很重要!!!)

$$X=\begin{pmatrix} x_1^T \\ \vdots \\ x_N^T \end{pmatrix}_{N\times n} \qquad X^T=(x_1,x_2,\cdots,x_N)_{n\times N}$$

$X^TX_{n\times n}$为样本的**协方差矩阵**

? X^T如何得到的，以及协方差矩阵如何计算让我们期待一下，到了分块矩阵再揭晓！