deepshare.net
深度之眼

机器学习—线性代数

导师：Johnson

deepshare.net
深度之眼

主要内容

1.相似矩阵的定义以及矩阵的对角化

2.矩阵对角化的条件以及对称矩阵的对角化

3.对角化在数据压缩算法中的应用

4.二次型以及矩阵的正定性

5.矩阵的正定性在机器学习线性回归算法中的运用(高级)

相似矩阵的定义以及矩阵的对角化

deepshare.net
深度之眼

Similar matrix

定义7 设A, B都是n阶矩阵，若有可逆矩阵P，使

$$P^{-1}AP = B,$$

则称B是A的相似矩阵，或说矩阵A与B相似.对A进行运算$P^{-1}AP$称为对A进行相似变换，可逆矩阵P称为把A变成B的相似变换矩阵.

定理3 若n阶矩阵A与B相似，则A与B的特征多项式相同，从而A与B的特征值亦相同.

推论 若n阶矩阵A与对角阵

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

相似，则$\lambda_1, \lambda_2, \cdots, \lambda_n$即是$A$的$n$个特征值.

相似矩阵的定义以及矩阵的对角化

下面我们要讨论的主要问题是：对n阶矩阵A，寻求相似变换矩阵P，使$P^{-1}AP = \Lambda$为对角阵，这就称为把矩阵A对角化.

假设已经找到可逆矩阵P，使$P^{-1}AP = \Lambda$为对角阵，我们来讨论P应满足什么关系.

把P用其列向量表示为

$$P = (p_1, p_2, \cdots, p_n)$$

由$P^{-1}AP = \Lambda$，得$AP = P\Lambda$，即

$$A(p_1, p_2, \cdots, p_n) = (p_1, p_2, \cdots, p_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

$$= (\lambda_1 p_1, \lambda_2 p_2, \cdots, \lambda_n p_n),$$

于是有 $$Ap_i = \lambda_i p_i \quad (i = 1,2, \cdots, n).$$

“==>”没有问题

“<==”需要P是可逆的

矩阵对角化的条件以及对称矩阵的对角化

deepshare.net 深度之眼

一般矩阵对角化的条件

定理4 n阶矩阵A与对角阵相似（即A能对角化）的充分必要条件是A有n个线性无关的特征向量.

定理2 设$\lambda_1, \lambda_2, \cdots, \lambda_m$是方阵$A$的$m$个特征值，$p_1, p_2, \cdots, p_m$依次是与之对应的特征向量，如果$\lambda_1, \lambda_2, \cdots, \lambda_m$各不相等，$p_1, p_2, \cdots, p_m$线性无关.

推论 如果n阶矩阵A的n个特征值互不相等，则A与对角阵相似.

矩阵对角化的条件以及对称矩阵的对角化

deepshare.net
深度之眼

对称矩阵对角化

定理5 对称阵的特征值为实数.

定理6 设λ_1,λ_2是对称阵A的两个特征值，p_1,p_2是对应的特征向量. 若$\lambda_1 \neq \lambda_2$，则p_1与p_2正交.

证 $\lambda_1 p_1 = Ap_1$，$\lambda_2 p_2 = Ap_2$，$\lambda_1 \neq \lambda_2$.

因A对称，故$\lambda_1 p_1^T = (\lambda_1 p_1)^T = (Ap_1)^T = p_1^T A^T = p_1^T A$，于是

$$\lambda_1 p_1^T p_2 = p_1^T A p_2 = p_1^T(\lambda_2 p_2) = \lambda_2 p_1^T p_2,$$

即

$$(\lambda_1 - \lambda_2) p_1^T p_2 = 0.$$

但$\lambda_1 \neq \lambda_2$，故$p_1^T p_2 = 0$，即p_1与p_2正交.

定理7 设A为n阶对称阵，则必有正交阵P，使$P^{-1}AP = P^T AP = \Lambda$，其中$\Lambda$是以$A$的$n$个特征值为对角元的对角阵.

矩阵对角化的条件以及对称矩阵的对角化

deepshare.net 深度之眼

对称矩阵对角化

推论 设A为n阶对称阵，λ是A的特征方程的k重根，则矩阵$A-\lambda E$的秩$R(A-\lambda E)=n-k$，从而对应特征值λ恰有k个线性无关的特征向量.

(i)求出A的全部互不相等的特征值$\lambda_1,\cdots,\lambda_s$，它们的重数依次为$k_1,\cdots,k_s(k_1+\cdots+k_s=n)$.

(ii)对每个k_i重特征值λ_i，求方程$(A-\lambda_i E)x=0$的基础解系，得k_i个线性无关的特征向量.再把它们正交化、单位化，得k_i个两两正交的单位特征向量.因$k_1+\cdots+k_s=n$，故总共可得n个两两正交的单位特征向量.

(iii)把这n个两两正交的单位特征向量构成正交阵P，便有$P^{-1}AP=P^TAP=\Lambda$.注意Λ中对角元的排列次序应与P中列向量的排列次序相对应.

矩阵对角化的条件以及对称矩阵的对角化

deepshare.net 深度之眼

例12 设 $A=\begin{pmatrix}0 & -1 & 1\\ -1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}$,

求一个正交阵P，使$P^{-1}AP=\Lambda$为对角阵.

解 由

$$|A-\lambda E|=\begin{vmatrix}-\lambda & -1 & 1\\ -1 & -\lambda & 1\\ 1 & 1 & -\lambda\end{vmatrix}=\begin{vmatrix}1-\lambda & \lambda-1 & 0\\ -1 & -\lambda & 1\\ 1 & 1 & -\lambda\end{vmatrix}=\begin{vmatrix}1-\lambda & 0 & 0\\ -1 & -1-\lambda & 1\\ 1 & 2 & -\lambda\end{vmatrix}$$
$$=(1-\lambda)(\lambda^2+\lambda-2)=-(\lambda-1)^2(\lambda+2),$$

求得A的特征值为$\lambda_1=-2,\lambda_2=\lambda_3=1.$

对应$\lambda_1=-2$，解方程$(A+2E)x=0$，由

$$A+2E=\begin{pmatrix}2 & -1 & 1\\ -1 & 2 & 1\\ 1 & 1 & 2\end{pmatrix}\sim\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix},$$

得基础解系$\xi_1=\begin{pmatrix}-1\\ -1\\ 1\end{pmatrix}$.将$\xi_1$单位化，得$p_1=\frac{1}{\sqrt{3}}\begin{pmatrix}-1\\ -1\\ 1\end{pmatrix}$.

矩阵对角化的条件以及对称矩阵的对角化

对应$\lambda_2 = \lambda_3 = 1$，解方程$(A - E)x = 0$，由

$$A - E = \begin{pmatrix} -1 & -1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

得基础解系$\xi_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$，$\xi_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

将ξ_2,ξ_3正交化：取$\eta_2 = \xi_2$，

$$\eta_3 = \xi_3 - \frac{[\eta_2, \xi_3]}{\|\eta_2\|^2}\eta_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

再将η_2,η_3单位化，得$p_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$，$p_3 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

矩阵对角化的条件以及对称矩阵的对角化

deepshare.net
深度之眼

将p_1, p_2, p_3构成正交矩阵

$$P = (p_1, p_2, p_3) = \begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix},$$

有 $$P^{-1}AP = P^TAP = \Lambda = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

对角化在数据压缩算法中的应用

A为n阶对称阵

$$A = P^{-1}\Lambda P = P^T\Lambda P$$

令$P^T = (p_1\ p_2\ \cdots\ p_n)$

$$A = (p_1\ p_2\ \cdots\ p_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}\begin{pmatrix}p_1^T \\ p_2^T \\ \vdots \\ p_n^T\end{pmatrix}$$

$$= \lambda_1 p_1 p_1^T + \lambda_2 p_2 p_2^T + \cdots + \lambda_n p_n p_n^T$$

二次型以及矩阵的正定性

deepshare.net
深度之眼

在解析几何中，为了便于研究二次曲线

$$ax^2 + bxy + cy^2 = 1$$

的几何性质，可以选择适当的坐标旋转变换

$$\begin{cases} x = x'\cos\theta - y'\sin\theta, \\ y = x'\sin\theta + y'\cos\theta, \end{cases}$$

把方程化为标准形

$$mx'^2 + ny'^2 = 1.$$

定义8 含有n个变量$x_1, x_2, \cdots, x_n$的二次齐次函数

$$f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2$$
$$+2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{n-1,n}x_{n-1}x_n$$

称为二次型.

二次型以及矩阵的正定性

deepshare.net
深度之眼

对于二次型，我们讨论的主要问题是：寻找可逆的线性变换

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n, \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots + c_{1n}y_n, \\ \cdots\ \cdots\ \cdots\ \cdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n \end{cases}$$

使二次型只含平方项，也就是用(7)代入(5)，能使

$$f = k_1y_1^2 + k_2y_2^2 + \cdots + k_ny_n^2,$$

这种只含平方项的二次型，称为二次型的标准形（或法式）.

如果标准形的系数$k_1, k_2, \cdots, k_n$只在$1, -1, 0$三个数中取值，能使

$$f = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2,$$

则称上式为二次型的规范形.

$$
\begin{aligned}
f &= a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\
&\quad + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\
&\quad + \cdots + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \\
&= \sum_{i,j=1}^{n} a_{ij}x_ix_j \\
&= (x_1, x_2, \cdots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}
\end{aligned}
$$

记

$$
A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},
$$

则二次型可记作

$$
f = x^T A x,
$$

其中A为对称阵.

例如，二次型$f = x^2 - 3z^2 - 4xy + yz$用矩阵记号写出来，就是

$$
f = (x, y, z) \begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.
$$

记

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

则二次型可记作

$$f=x^T Ax,$$

其中A为对称阵.

如果A是对角矩阵该多棒，一下子就是标准型甚至规范型了

推论 对称阵A为正定的充分必要条件是：A的特征值全为正.

定义10 设有二次型$f(x)=x^TAx$，如果对任何$x\neq 0$，都有$f(x)>0$(显然$f(0)=0$),则称f为正定二次型，并称对称阵A是正定的；如果对任何$x\neq 0$，都有$f(x)<0$，则称f为负定二次型，并称对称阵A是负定的.

定理10 n元二次型$f=x^TAx$为正定的充分必要条件是：它的标准形的n个系数全为正，即它的规范形的n个系数全为1，亦即它的正惯性指数等于n.

补充半正定知识点

矩阵的正定性在机器学习线性回归算法中的运用(高级)

deepshare.net
深度之眼

$x_1, x_2, \cdots, x_N, x_i \in \mathbb{R}^n$

$y_1, y_2, \cdots, y_N, y_i \in \mathbb{R}^1$

$y_1 = x_{11}a_1 + x_{12}a_2 + \cdots + x_{1n}a_n$

$y_2 = x_{21}a_1 + x_{22}a_2 + \cdots + x_{2n}a_n$

$\vdots$

$y_N = x_{N1}a_1 + x_{N2}a_2 + \cdots + x_{Nn}a_n$

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Nn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}$$

$X_{N\times n}a_{n\times 1} = Y_{N\times 1}$

当$N = n$且$X_{N\times n}$可逆时:

$a = X^{-1}Y$

一般情况：$N \neq \mathrm{n}$

$\min\|xa - Y\|^2 = J \qquad \frac{\partial J}{\partial a} = x^T(xa - Y) = 0$

$x^T xa = x^T Y \qquad x^T x$是否可逆?

1. $\mathrm{N} > \mathrm{n}$

如$N = 5, n = 3 \quad (x^T x)_{3\times 3}$一般是可逆的

$a = \underline{(x^T x)^{-1} x^T} Y$

伪逆

2. $N < n$

如$N = 3, n = 5 \quad (x^T x)_{5\times 5}$

$R(x^T x) \leq R(x) \leq 3$

故$x^T x$不可逆

补充：$R(AB) \ll R(A)$ or$R(B)$

矩阵的正定性在机器学习线性回归算法中的运用(高级)

deepshare.net 深度之眼

此刻 $J=\|xa-Y\|^2+\lambda\|a\|^2$

$$\frac{\partial J}{\partial a}=x^Tx a-x^TY+\lambda a=0$$

$$(x^Tx+\lambda I)a=x^TY$$

必为可逆

$$a=(x^Tx+\lambda I)^{-1}x^TY$$

x^Tx为对称矩阵可对角化

$$x^Tx=\mathrm{p}^{-1}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\mathrm{p}$$

$$|x^Tx|=\lambda_1\lambda_2\cdots\lambda_{\mathrm{n}}$$

1. $a^T(x^Tx)a=(xa)^T(xa)\geq 0\ \rightarrow\ \lambda_i\geq 0$

   $|x^Tx|$仍可能为0，不一定可逆

2. $a^T(x^Tx+\lambda I)a=(xa)^T(xa)+\lambda a^Ta>0\ \rightarrow\ \lambda_i>0$

   $|x^Tx+\lambda I|>0$恒成立，一定可逆

deepshare.net
深度之眼

联系我们：

电话：18001992849

邮箱：service@deepshare.net

Q Q： 2677693114

公众号

客服微信