大数据的矩阵计算基础——第9周

法律声明

【声明】本视频和幻灯片为炼数成金网络课程的教学资料，所有资料只能在课程内使用，不得在课程以外范围散播，违者将可能被追究法律和经济责任。

关注炼数成金企业微信

■提供全面的数据价值资讯，涵盖商业智能与数据分析、大数据、企业信息化、数字化技术等，各种高性价比课程信息，赶紧掏出您的手机关注吧！

本周内容

- 二次型的条件优化
- 矩阵分解
    - 谱分解
    - Jordan分解
    - SVD分解
    - 其他分解

条件优化

- 单位向量的等价形式

- 限定在$x^T x = 1$条件下二次型Q(x)的最大值与最小值

求 $Q(\boldsymbol{x})=9x_1^2+4x_2^2+3x_3^2$，在限制条件 $\boldsymbol{x}^T\boldsymbol{x}=1$ 下的最大值和最小值.

条件优化

- 令A=$\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$，求限定在$x^T x = 1$条件下二次型z=Q(x)的最大值与最小值

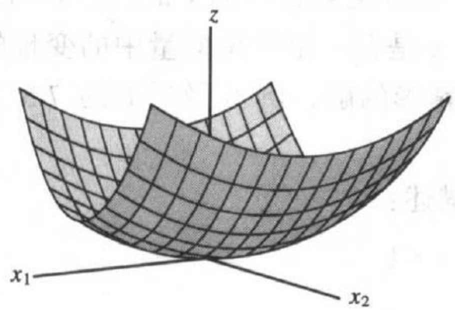

图 7-8　$z=3x_1^2+7x_2^2$

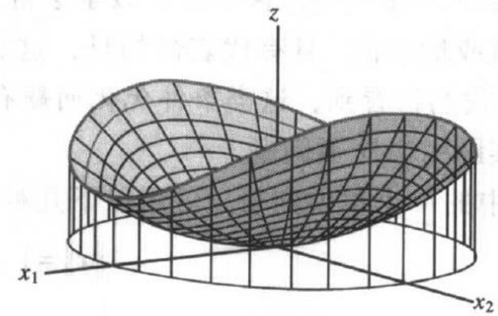

图 7-9　$z=3x_1^2+7x_2^2$ 和圆柱 $x_1^2+x_2^2=1$ 的交线

$$m=\min\{x^{T}Ax:\|x\|=1\}\ ,\ M=\max\{x^{T}Ax:\|x\|=1\} \qquad (2)$$

定理 6 设 A 是对称矩阵，且 m 和 M 的定义如（2）式所示，那么 M 是 A 的最大特征值 λ_1，m 是 A 的最小特征值，如果 x 是对应 M 的单位特征向量 u_1，那么 $x^{T}Ax$ 的值等于 M，如果 x 是对应 m 的单位特征向量，$x^{T}Ax$ 的值等于 m.

例 3 令 $A=\begin{bmatrix}3&2&1\\2&3&1\\1&1&4\end{bmatrix}$，求二次型 $x^{T}Ax$ 在限制条件 $x^{T}x=1$ 下的最大值，且求一个可以取到该最大值的单位向量.

条件优化

DATAGURU 炼数成金

定理 7 设 A, λ_1 和 u_1 如定理 6 所示. 在如下条件限制下

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1,\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_1=0$$

$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的最大值是第二大特征值 λ_2，且这个最大值，可以在 $\boldsymbol{x}$ 是对应 λ_2 的特征向量 $\boldsymbol{u}_2$ 处达到.

例 4 求 $9x_1^2+4x_2^2+3x_3^2$ 的最大值，其限制条件是 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1$ 和 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_1=0$，这里 $\boldsymbol{u}_1=(1,0,0)$. 注意到 $\boldsymbol{u}_1$ 是二次型对应的矩阵的最大特征值 $\lambda=9$ 对应的单位特征向量.

例 5 令 A 表示例 3 中的矩阵，且 $\boldsymbol{u}_1$ 是对应矩阵 A 最大特征值的特征向量，求 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的最大值，其限制条件是

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1,\quad \boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_1=0 \tag{4}$$

条件优化

定理 8 设 A 是一个 $n\times n$ 对称矩阵，且其正交对角化为 $A=PDP^{-1}$，将对角矩阵 D 上的元素重新排列，使得 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$，且 P 的列是其对应的单位特征向量 $u_1,\cdots,u_n$. 那么对 $k=2,\cdots,n$ 时，在以下限制条件下

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=1,\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_1=0,\ \cdots,\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u}_{k-1}=0$$

$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ 的最大值是特征值 λ_k，且这个最大值在 $\boldsymbol{x}=\boldsymbol{u}_k$ 处可以达到.

条件优化

例 6　在下一年度，一个县政府计划修 x 百英里的公路和桥梁，并且修整 y 百英亩的公园和娱乐场所，政府部门必须确定在两个项目上如何分配它的资源（资金、设备和劳动等等），如果更划算的话，可以同时开始两个项目，而不是仅开始一个项目，那么 x 和 y 必须满足下面限制条件

$$4x^2+9y^2\leqslant 36$$

见图 7-10，每个阴影可行集中的点 (x, y)，表示一个可能的该年度公共工作计划．在限制曲线 $4x^2+9y^2=36$ 上的点，使资源利用达到最大可能．

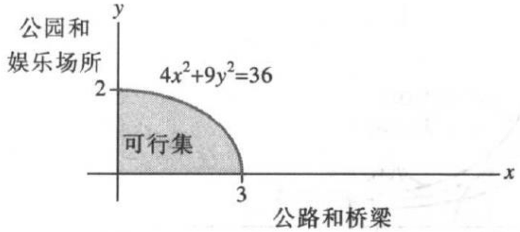

图 7-10　公共工作计划

为选择它的公共计划，县政府需要考虑居民的意见，为度量居民分配各类工作计划 (x, y) 的值或效用，经济学家有时利用下面的函数

$$q(x, y)=xy$$

其中使 $q(x, y)$ 为常数的 (x, y) 点的集合，称为无差别曲线，从图 7-11 中可以看到三条这样的曲线，沿着无差别曲线的点对应的选择，表示居民作为一个群体有相同的价值观[㊀]．求公共工作计划，使得效用函数 q 最大．

矩阵分解

- 矩阵的分解：
  - $A = A_1 + A_2 + \cdots \ldots + A_n$
  - $A = A_1 A_2 \ldots \ldots A_n$
- 矩阵分解的常见方法
- 矩阵分解的意义

对角化

- 从线性变换的角度：若线性变换f关于线性空间V的某个基的矩阵为对角矩阵，则称f可对角化
- 从矩阵的角度：若n阶方阵A相似与对角矩阵D，即存在可逆矩阵P，使得$A = PDP^{-1}$,则称A可对角化
- 矩阵A可对角化的充要条件是:A有n个线性无关的特征向量
- 有n个不同的特征值的矩阵A可对角化（充分条件而不是必要条件）
- 若$A = PDP^{T}$，其中P为正交矩阵，D为对角矩阵，则称A可正交对角化

谱分解

- 假设n阶方阵A可以正交对角化，即$A = PDP^T$，其中P的列向量为矩阵A的单位正交特征向量$u_1, u_2, \ldots \ldots, u_n$,D为由A的特征值构成的对角矩阵，则

$$A = PDP^T = \lambda_1 u_1 u_1^T + \cdots \ldots + \lambda_n u_n u_n^T$$

- 称矩阵A这样表示为A的谱分解

谱分解

构造矩阵 A 的一个谱分解，已知 A 有以下正交对角化分解.

$$A=\begin{bmatrix}7 & 2\\ 2 & 4\end{bmatrix}=\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}\begin{bmatrix}8 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$$

Jordan标准型

- Jordan块

$$\begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & \ddots & 0 \\ 0 & 0 & \ddots & 1 \\ 0 & 0 & \cdots & \lambda_i \end{bmatrix} \in F^{r_i \times r_i}$$

- Jordan矩阵

$$\begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix}$$

Jordan标准型

- Jordan分解
- 设A是n阶方阵，则存在n阶可逆矩阵T使得

$$A = TJT^{-1}$$

其中$J = diag(J_{n_1}(\lambda_1), J_{n_2}(\lambda_2), \ldots\ldots, J_{n_k}(\lambda_k))$,称J为A的Jordan标准型，T为变换矩阵。

- 称$A = TJT^{-1}$为A的Jordan分解

Jordan标准型

- J的特征值$\lambda_1, \lambda_2, \ldots\ldots \lambda_k$是A的特征值
- λ_i的几何重数与代数重数

$$\begin{bmatrix} 2 & 0 & & \\ & 2 & 1 & \\ & & 2 & 0 \\ & & & 1 \end{bmatrix}$$

Jordan标准型

- 将一般方阵化为Jordan标准型
- 特征向量法

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

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

Jordan标准型

- 变换矩阵T的求解
- $\exists k$阶方阵T，使得$A = TJT^{-1}$， $\Longrightarrow AT = TJ$

$$T = (T_1, T_2, \ldots\ldots, T_k)$$

$$T_i = (t_1^i, t_2^i, \ldots\ldots, t_{n_i}^i)$$

- Jordan链

$$\begin{cases} At_1^i = \lambda_i t_1^i \\ At_2^i = \lambda_i t_2^i + t_1^i \\ \ldots \\ At_{n_i}^i = \lambda_i t_{n_i}^i + t_{n_{i-1}}^i \end{cases}$$

Jordan标准型

- 例子

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

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

奇异值分解（SVD分解）

例 1　如果 $A=\begin{bmatrix}4 & 11 & 14\\ 8 & 7 & -2\end{bmatrix}$，那么线性变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$，将 $\mathbb{R}^3$ 中的单位球 $\{\boldsymbol{x}:\|\boldsymbol{x}\|=1\}$ 映射为 $\mathbb{R}^2$ 中的椭圆，见图 7-13，找出使得长度 $\|A\boldsymbol{x}\|$ 最大的一个单位向量 $\boldsymbol{x}$，且计算这个最大长度.

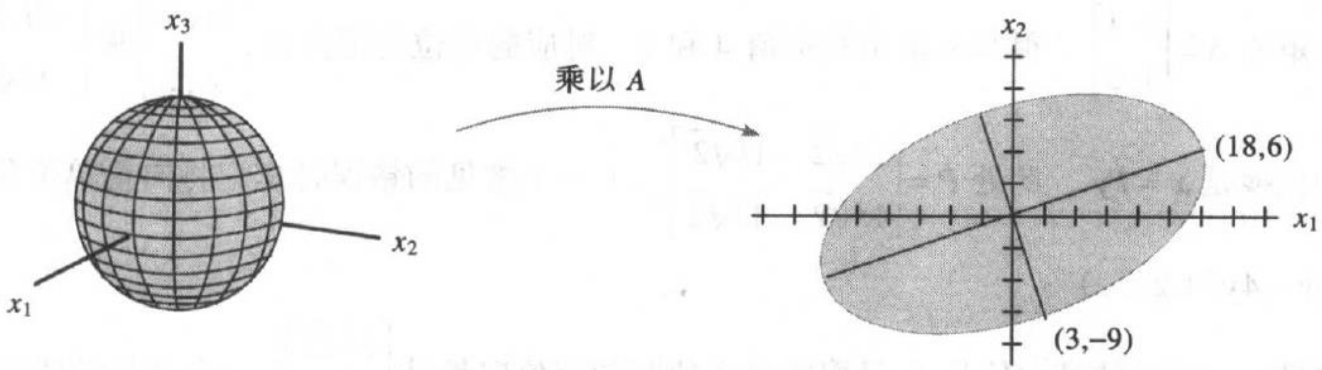

图 7-13　从 $\mathbb{R}^3$ 到 $\mathbb{R}^2$ 的一个变换

奇异值分解（SVD分解）

- 设A是m*n的矩阵
- 奇异值
- 例：求矩阵A的奇异值：$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$

奇异值分解（SVD分解）

- 若$\{v_1, \ldots \ldots, v_n\}$是包含$A^T A$特征向量的$R^n$上的单位正交基，对应$A^T A$的特征值$\lambda_1 \geq \lambda_2 \geq \cdots \ldots \geq \lambda_n$.若A有r个非零奇异值，那么$\{Av_1, \ldots \ldots, Av_n\}$是ColA的一个正交基，且rankA=r。

奇异值分解（SVD分解）

- $m \times n$对角矩阵$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$
- $m \times m$正交矩阵U——左奇异向量
- $n \times n$正交矩阵V——右奇异向量
- A的一个奇异值分解：$A = U\Sigma V^T$

奇异值分解（SVD分解）

- 奇异值分解三部曲：
- 1. 将矩阵$A^T A$正交对角化
- 2. 算出V和Σ
- 3. 构造U

奇异值分解（SVD分解）

- 例：求 $A=\begin{bmatrix}4 & 11 & 14\\ 8 & 7 & -2\end{bmatrix}$ 的一个奇异值分解.

求 $A=\begin{bmatrix}1 & -1\\ -2 & 2\\ 2 & -2\end{bmatrix}$ 的一个奇异值分解.

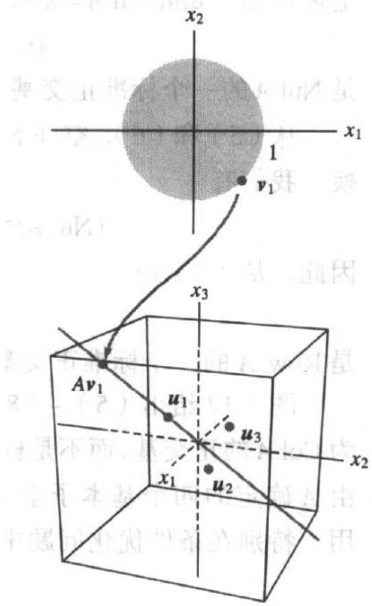

图 7-15

奇异值分解的应用

- 计算存储图形——将图形分解成象素（pixels）的一个矩形的数阵，其中的信息就可以用一个矩阵A=（a_{ij}）$_{m\times n}$来存储。矩阵A的元素a_{ij}是一个正的数，它相应于象素的灰度水平（gray level）的度量值。
- 由于一般来讲，相邻的象素会产生相近的灰度水平值，因此有可能在满足图像清晰度要求的条件下，将存储一个m×n阶矩阵需要存储的m×n个数减少到n+m+1的一个倍数。
- 原矩阵$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$
- 压缩矩阵 $A_k = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_k u_k v_k^T, k \le r$
- 应用实例：https://yihui.shinyapps.io/imgsvd/

LU分解

- 对于n阶方阵A，存在下三角矩阵L，上三角矩阵U，使得A=LU。称矩阵A这样的表示为A的LU分解

- 如：

$$A=\begin{bmatrix}1&0&7\\3&2&0\\1&1&1\end{bmatrix}=\begin{pmatrix}1&0&0\\3&1&0\\1&\frac{1}{2}&1\end{pmatrix}\begin{pmatrix}1&0&7\\0&2&-21\\0&0&\frac{9}{2}\end{pmatrix}=LU$$

LU分解的应用

DATAGURU 炼数成金

- 简化求解线性方程的算法

$$Ax = b \qquad LUx = b$$

消元过程 $Ly = b$

回代过程 $Ux = y$

- 例：解方程组AX=b，其中$A = \begin{pmatrix} 1 & 0 & 7 \\ 3 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}, b = \begin{pmatrix} -6 \\ 3 \\ 0 \end{pmatrix}$

其他分解

- 如果实非奇异矩阵A能够化成正交矩阵Q与实非奇异上三角矩阵R的乘积，即$A=QR$，则称为A的QR分解

- 设A是m*n阶矩阵，则存在m*r阶矩阵B，r*n阶矩阵C，使得A=BC。其中B为列满秩矩阵，C为行满秩矩阵。称此分解为A的满秩分解

炼数成金逆向收费式网络课程

- Dataguru（炼数成金）是专业数据分析网站，提供教育，媒体，内容，社区，出版，数据分析业务等服务。我们的课程采用新兴的互联网教育形式，独创地发展了逆向收费式网络培训课程模式。既继承传统教育重学习氛围，重竞争压力的特点，同时又发挥互联网的威力打破时空限制，把天南地北志同道合的朋友组织在一起交流学习，使到原先孤立的学习个体组合成有组织的探索力量。并且把原先动辄成千上万的学习成本，直线下降至百元范围，造福大众。我们的目标是：低成本传播高价值知识，构架中国第一的网上知识流转阵地。
- 关于逆向收费式网络的详情，请看我们的培训网站 http://edu.dataguru.cn

Thanks

FAQ时间