

阅读 ▶ 最新

知识

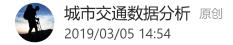
产业

深度

专栏

SOTA

产业



浅谈「正定矩阵」和「半正定矩阵」

在众多的机器学习模型中,线性代数的身影无处不在,当然,我们也会时常碰到线性代数中的正定矩阵和半正定矩阵。例如,多元正态分布的协方差矩阵要求是半正定的。

1. 基本的定义

正定和半正定这两个词的英文分别是positive definite和positive semi-definite,其中,definite是一个形容词,表示"明确的、确定的"等意思。

初学线性代数的读者可能会被这两个词"唬住",但正定矩阵和半正定矩阵的定义实际上是很简单的(不考虑复数构成的矩阵):

【定义1】给定一个大小为 $m{n} imes m{n}$ 的实对称矩阵 $m{A}$,若对于任意长度为 $m{n}$ 的非零向量 $m{x}^T A m{x} > m{0}$ 恒成立,则矩阵 $m{A}$ 是一个正定矩阵。

【例1】单位矩阵 $I \in \mathbb{R}^{2 \times 2}$ 是否是正定矩阵?

由于 $oldsymbol{x}
eq oldsymbol{0}$,故 $oldsymbol{x}^T I oldsymbol{x} > oldsymbol{0}$ 恒成立,即单位矩阵 $oldsymbol{I} \in \mathbb{R}^{2 imes 2}$ 是正定矩阵。

单位矩阵是正定矩阵 (positive definite)。



知识

产业

深度 专栏 SOTA

产业对

$$oldsymbol{x}^T I oldsymbol{x} = oldsymbol{x}^T oldsymbol{x}$$
 $= x_1^2 + x_2^2 + \dots + x_n^2 > 0$

$$A = egin{bmatrix} 2 & -1 & 0 \ -1 & 2 & -1 \ 0 & -1 & 2 \end{bmatrix} \in \mathbb{R}^{3 imes 3}$$
 是否是正定矩阵?

【例2】 实对称矩阵

$$oldsymbol{x} = egin{bmatrix} x_1 \ x_2 \ x_3 \end{bmatrix} \in \mathbb{R}^3$$
 هرين هيرين المحمدة المحم

$$oldsymbol{x}^T A oldsymbol{x} = \left[egin{array}{ccc} (2x_1 - x_2) & (-x_1 + 2x_2 - x_3) & -x_2 + 2x_3 \end{array}
ight] egin{bmatrix} x_1 \ x_2 \ x_3 \end{array}$$

$$=x_1^2+(x_1-x_2)^2+(x_2-x_3)^2+x_3^2>0$$

因此, 矩阵 A 是正定矩阵。

【定义2】给定一个大小为 $oldsymbol{n} imesoldsymbol{n}$ 的实对称矩阵 $oldsymbol{A}$,若对于任意长度为 $oldsymbol{n}$ 的向量 $oldsymbol{x}$, a $oldsymbol{x}^T A oldsymbol{x} \geq 0$ leads, where $oldsymbol{A}$ leads $oldsymbol{A}$ leads $oldsymbol{A}$ leads $oldsymbol{x}$

根据正定矩阵和半正定矩阵的定义,我们也会发现:半正定矩阵包括了正定矩阵,与非负实数 (non-negative real number)和正实数 (positive real number)之间的关系很像。



图1 正实数与负实数,图片来源于https://en.wikipedia.org/wiki/Real_number



▶ 最新 知识 产业 深度 专栏

SOTA

产业

在初中数学中,我们学习了二次函数 $oldsymbol{y}=ax^2$,该函数的曲线会经过坐标原点,当参数 a>0 时,曲线的"开口"向上,参数 a<0 时,曲线的"开口"向下。

以 $y=2x^2$ 为例,曲线如下:

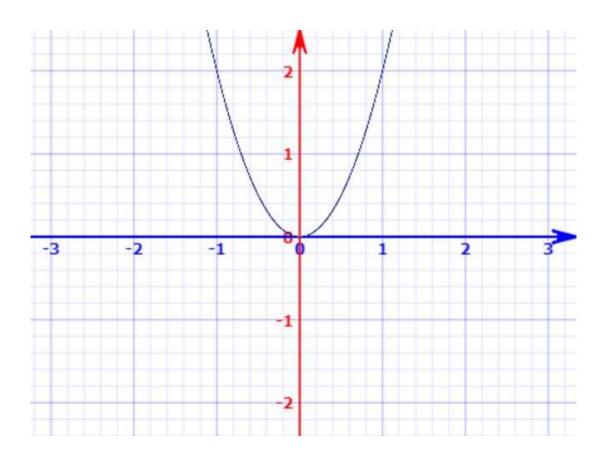


图2 二次函数曲线

实际上,我们可以将 $y = \boldsymbol{x}^T A \boldsymbol{x}$ 视作 $y = a x^2$ 的多维表达式。

 $_{ ext{当我们希望}}\; oldsymbol{y} = oldsymbol{x}^T A oldsymbol{x} \geq 0$ 对于任意向量 $oldsymbol{x}$ 都恒成立,就要求矩阵 $oldsymbol{A}$ 是一个半正 定矩阵,对应于二次函数, $y=ax^2>0, \forall x$ 需要使得 $a\geq 0$.

 $_{
m Sh,\ A}$ $_{
m C}$ $_{
m Ch}$ $_{
m$ 成立。



知识 产业 深度 专栏

SOTA

产业

3. 正定矩阵和半正定矩阵的直观解释

若给定任意一个正定矩阵 $A \in \mathbb{R}^{n \times n}$ 和一个非零向量 $\boldsymbol{x} \in \mathbb{R}^n$,则两者相乘得到的 $egin{aligned} egin{aligned} egin{aligned} oldsymbol{y} & oldsymbol{x} & oldsymbol{x} \end{aligned} & oldsymbol{y} & oldsymbol{y} & oldsymbol{z} \end{aligned} & oldsymbol{y} & oldsymbol{y} & oldsymbol{z} \end{aligned} & oldsymbol{z} & oldsymbol{x} & oldsymbol{x}^T A oldsymbol{x} > oldsymbol{0} \end{aligned}$

$$m{x} = egin{bmatrix} 2 \ 1 \end{bmatrix} \in \mathbb{R}^2$$
 ,对于单位矩阵 $I = egin{bmatrix} 1 & 0 \ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2 imes 2}$,则

$$oldsymbol{y} = Ioldsymbol{x} = oldsymbol{x} = egin{bmatrix} 2 \ 1 \end{bmatrix}$$

 $_{oldsymbol{eta}}$ $oldsymbol{x},oldsymbol{y}\in\mathbb{R}^2$ 之间的夹角为

$$\cos\langle oldsymbol{x}, oldsymbol{y}
angle = rac{oldsymbol{x}^Toldsymbol{y}}{||oldsymbol{x}||\cdot||oldsymbol{y}||}$$

$$=rac{2 imes 2+1 imes 1}{\sqrt{2^2+1^2}\cdot\sqrt{2^2+1^2}}$$

= 1

即两个向量之间的夹角为0°.

$$oldsymbol{y} = Aoldsymbol{x} = egin{bmatrix} 0 \ 2 \ 0 \end{bmatrix}$$



知识 产业

专栏 深度

SOTA

产业

$$\langle \cos \langle oldsymbol{x}, oldsymbol{y}
angle = rac{oldsymbol{x}^T oldsymbol{y}}{||oldsymbol{x}|| \cdot ||oldsymbol{y}||} = rac{\sqrt{6}}{3}$$

即两个向量之间的夹角小 \mp $\frac{-}{2}$

若给定任意一个正定矩阵 $A \in \mathbb{R}^{n \times n}$ 和一个向量 $\boldsymbol{x} \in \mathbb{R}^n$,则两者相乘得到的向 $_{oxedge} \; oldsymbol{y} = Aoldsymbol{x} \in \mathbb{R}^n \;$ 与向 $oxedge} \; oldsymbol{x} \;$ 的夹角恒小于或等于 $\; oldsymbol{\overline{2}} \;$ 、(等价于: $\; oldsymbol{x}^T Aoldsymbol{x} \geq oldsymbol{0} \;$)

3. 为什么协方差矩阵要是半正定的?

在概率论与数理统计中,我们都学习的协方差矩阵的定义:

对于任意多元随机变量 $oldsymbol{t}$,协方差矩阵为 $oldsymbol{C} = \mathbb{E}\left[(oldsymbol{t} - ar{oldsymbol{t}})^T
ight]$

现给定任意一个向量 化 ,则

$$oldsymbol{x}^T C oldsymbol{x} = oldsymbol{x}^T \mathbb{E}\left[(oldsymbol{t} - ar{oldsymbol{t}})(oldsymbol{t} - ar{oldsymbol{t}})^T
ight] oldsymbol{x}$$

$$oxed{\mathbf{z}} = \mathbb{E}\left[oldsymbol{x}^T(oldsymbol{t} - ar{oldsymbol{t}})(oldsymbol{t} - ar{oldsymbol{t}})^Toldsymbol{x}
ight]$$

$$=\mathbb{E}(s^2)=\sigma_s^2$$

其中,

$$\sigma_s = oldsymbol{x}^T (oldsymbol{t} - ar{oldsymbol{t}}) = (oldsymbol{t} - ar{oldsymbol{t}})^T oldsymbol{x}$$

 $_{ ext{h+}} \sigma_s^2 \geq 0$,因此, $oldsymbol{x}^T C oldsymbol{x} \geq 0$, $_{ ext{h} ext{b} ext{b} ext{b} ext{E}}$ C $_{ ext{E} ext{E} ext{E} ext{E} ext{D} ext{c}}$.

相关参考:



阅读 🕨 最新

知识

产业

深度

专栏

SOTA

产业对



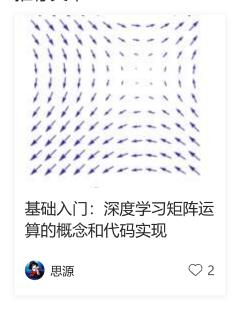
城市交通数据分析

Machine learning for urban traffic data analytics.

入门

♥4 □2 %

推荐文章











Chris

 \square \bigcirc

你有一个错误, 在举矩阵的例子时, 错将半正定写成了正定, 因为大于等于零是半正定的





阅读 ▶ 最新

知识

产业

深度

专栏

SOTA

产业对



关于我们

全球人工智能信息服务

友情链接: Synced Global

机器之心 Medium 博客

PaperWeekly 动脉网 艾耕科技

%





©2020 机器之心 (北京) 科技有限公司 京 ICP 备 14017335号-2