



矩阵求导入门

0

覃雄派



矩阵求导入门

提纲



• 矩阵求导入门



• 矩阵求导参考资料

The Matrix Cookbook

[http://matrixcookbook.com]

Kaare Brandt Petersen Michael Syskind Pedersen

Version: November 15, 2012

这里先进行简单入门,然后直接使用该cook book的一些结论

https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf





π @ R U

矩阵求导



• 矩阵求导入门

$$\mathbf{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix}$$

按照"列"组织向量

$$\mathbf{x}^T \boldsymbol{\theta} = \begin{bmatrix} 1 & x_1 & x_2 & \dots & x_d \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix} = \theta_0 + x_1 \theta_1 + x_2 \theta_2 + \dots + x_d \theta_d$$

向量点乘

• 矩阵求导入门

$$\mathbf{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$oldsymbol{ heta} = egin{bmatrix} heta_0 \ heta_1 \ heta_2 \ heta_d \end{bmatrix}$$

$$\mathbf{x}^T \boldsymbol{\theta} = \begin{bmatrix} 1 & x_1 & x_2 & \dots & x_d \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix} = \theta_0 + x_1 \theta_1 + x_2 \theta_2 + \dots + x_d \theta_d$$

$$1. \theta_0 + x_1 \theta_1 + x_2 \theta_2 + \dots + x_d \theta_d$$
 对 θ_0 、 θ_1 、 θ_2 、 \dots 、 θ_d 求导

2.得到

$$1 \quad x_1 \quad x_2 \quad \dots \quad x_d$$

.按照列向量来组织为 x₂ : x₄

• 矩阵求导入门

$$\mathbf{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$oldsymbol{ heta} = egin{bmatrix} heta_0 \ heta_1 \ heta_2 \ heta_d \end{bmatrix}$$

$$\mathbf{x}^T \boldsymbol{\theta} = \begin{bmatrix} 1 & x_1 & x_2 & \dots & x_d \end{bmatrix}$$

$$1. \theta_0 + x_1 \theta_1 + x_2 \theta_2 + \dots + x_d \theta_d$$
 对 θ_0 、 θ_1 、 θ_2 、 \dots 、 θ_d 求导

2.得到

$$1 \quad x_1 \quad x_2 \quad \dots \quad x_d$$

3.按照列向量来组织为



$$\frac{\partial \mathbf{x}^T \boldsymbol{\theta}}{\partial \boldsymbol{\theta}} = \frac{\partial \boldsymbol{\theta}^T \mathbf{x}}{\partial \boldsymbol{\theta}} = \mathbf{x}$$

第一个重要公式

 $=\theta_0 + x_1\theta_1 + x_2\theta_2 + \dots + x_d\theta_d$

• 矩阵求导入门

$$\mathbf{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$

$$oldsymbol{ heta} = egin{bmatrix} heta_0 \ heta_1 \ heta_2 \ heta_d \end{bmatrix}$$

$$\mathbf{x}^T \boldsymbol{\theta} = \begin{bmatrix} 1 & x_1 & x_2 & \dots & x_d \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix} = \theta_0 + x_1 \theta_1 + x_2 \theta_2 + \dots + x_d \theta_d$$

$$\frac{\partial \boldsymbol{a}^T}{\partial \boldsymbol{a}^T} = \frac{\partial \boldsymbol{a}^T$$

$$\theta_0+x_1\theta_1+x_2\theta_2+...+x_d\theta_d$$

对 θ_0 、 θ_1 、 θ_2 、...、 θ_d
求导得到

 $1 \quad x_1 \quad x_2 \quad \dots \quad x_d$

按照列向量来组织为 $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$

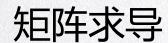
$$oldsymbol{x}$$
 列向量

$$\boldsymbol{a}$$
 列向量

$$rac{\partial oldsymbol{a}^T \mathbf{x}}{\partial \mathbf{x}} = rac{\partial \mathbf{x}^T oldsymbol{a}}{\partial \mathbf{x}} = oldsymbol{a}$$









- 矩阵求导入门
 - 列向量
 - a 列向量

$$rac{\partial oldsymbol{a}^T \mathbf{x}}{\partial \mathbf{x}} = rac{\partial \mathbf{x}^T oldsymbol{a}}{\partial \mathbf{x}} = oldsymbol{a}$$



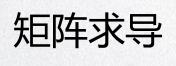
 $\mathbf{b}^T \mathbf{A}$ 看作一个整体

$$rac{\partial \mathbf{b}^T \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}^T \mathbf{b}$$



• 矩阵求导入门

- $oldsymbol{x}$ 列向量
- a 列向量



$$rac{\partial oldsymbol{a}^T \mathbf{x}}{\partial \mathbf{x}} = rac{\partial \mathbf{x}^T oldsymbol{a}}{\partial \mathbf{x}} = oldsymbol{a}$$



有两个x,分别 求导,累加



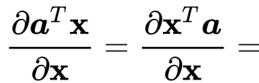
$$\mathbf{x}^T \mathbf{A} \mathbf{x}$$

$$rac{\partial \mathbf{x}^T \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = oldsymbol{A} oldsymbol{x} + oldsymbol{A}^T oldsymbol{x}$$



SHIVERS/7/LOG CHINA

- 矩阵求导入门
 - $oldsymbol{x}$ 列向量
 - $oldsymbol{a}$ 列向量



有两个x,分别 求导,累加







2 Ax 如果A是对称矩阵

$$\frac{\partial \mathbf{x}^T \mathbf{A} \mathbf{x}}{\partial \mathbf{x}}$$

$$=oldsymbol{A}oldsymbol{x}+oldsymbol{A}^Toldsymbol{x}$$



