矩阵求导入门

覃雄派

矩阵求导入门

提纲

- 矩阵求导入门

矩阵求导

- 矩阵求导参考资料

The Matrix Cookbook

[http://matrixcookbook.com]

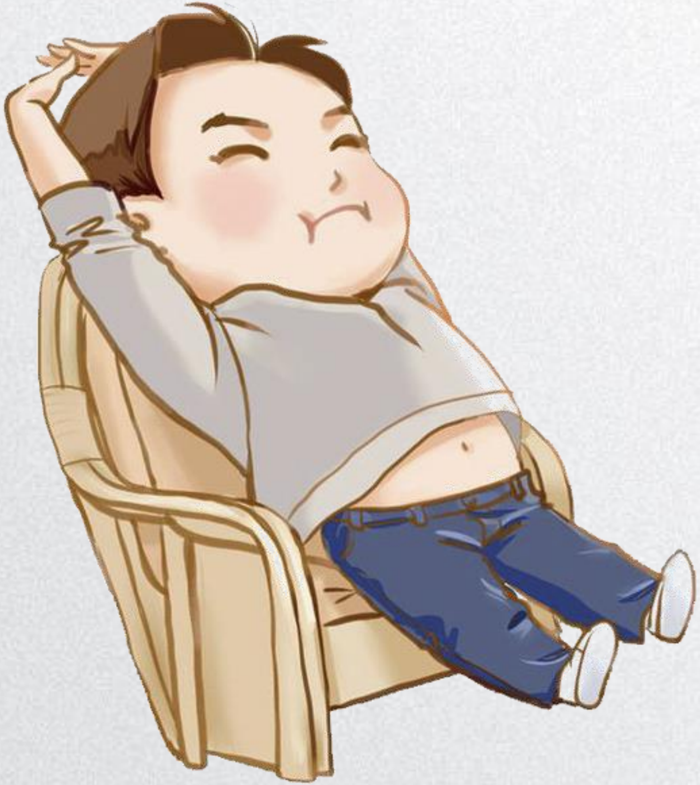
Kaare Brandt Petersen
Michael Syskind Pedersen

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

矩阵求导



矩阵求导

- 矩阵求导入门

按照“列”组织向量

$$\mathrm{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \qquad \boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix}$$

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

向量点乘

π@RUC

矩阵求导

- 矩阵求导入门

$$\mathrm{x}=\begin{bmatrix}1\\x_1\\x_2\\\vdots\\x_d\end{bmatrix}\qquad\boldsymbol{\theta}=\begin{bmatrix}\theta_0\\\theta_1\\\theta_2\\\vdots\\\theta_d\end{bmatrix}$$

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

1. $\theta_0+x_1\theta_1+x_2\theta_2+\ldots+x_d\theta_d$ 对 θ_0、θ_1、θ_2、…、θ_d 求导
2. 得到
$$1 \quad x_1 \quad x_2 \quad \ldots \quad x_d$$
3. 按照列向量来组织为 $\begin{bmatrix}1\\x_1\\x_2\\\vdots\\x_d\end{bmatrix}$

矩阵求导

- 矩阵求导入门

$$\mathrm{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \qquad \boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix}$$

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

1. $\theta_0+x_1\theta_1+x_2\theta_2+\ldots+x_d\theta_d$ 对θ_0、θ_1、θ_2、…、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 $\begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$

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

第一个重要公式

矩阵求导

- 矩阵求导入门

$$\mathrm{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \qquad \boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix}$$

$$\mathrm{x}^T\theta = [1 \quad x_1 \quad x_2 \quad \dots \quad x_d] \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_d \end{bmatrix} = \theta_0 + x_1\theta_1 + x_2\theta_2 + \dots + x_d\theta_d$$

$\theta_0 + x_1\theta_1 + x_2\theta_2 + \dots + x_d\theta_d$ 对θ_0、θ_1、θ_2、…、θ_d求导得到

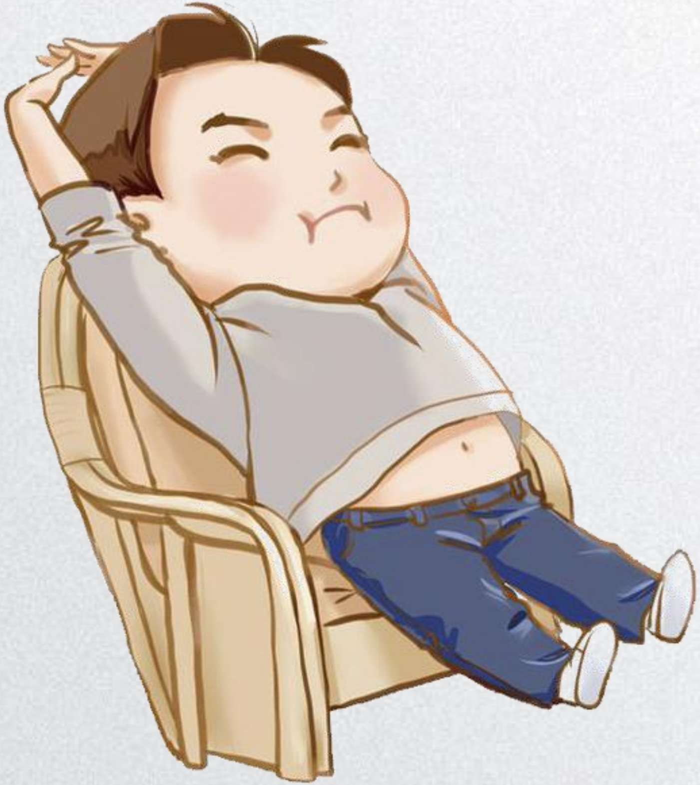
$1 \quad x_1 \quad x_2 \quad \dots \quad x_d$

按照列向量来组织为$\begin{bmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$

$\boldsymbol{x}$ 列向量
$\boldsymbol{a}$ 列向量

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

矩阵求导



π@RUC

矩阵求导

- 矩阵求导入门

$\boldsymbol{x}$ 列向量

$\boldsymbol{a}$ 列向量

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

$\mathbf{b}^T \mathbf{A}$ 看作一个整体

$$\boxed{\frac{\partial \mathbf{b}^T \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = \boldsymbol{A}^T \boldsymbol{b}}$$

π@RUC

矩阵求导

- 矩阵求导入门

$\boldsymbol{x}$	列向量
$\boldsymbol{a}$	列向量

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

有两个x，分别求导，累加

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

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

$$\boxed{\frac{\partial \mathbf{x}^T \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}^T\boldsymbol{x}}$$

π@RUC

矩阵求导

- 矩阵求导入门

$\boldsymbol{x}$	列向量
$\boldsymbol{a}$	列向量

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

有两个x，分别求导，累加

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

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

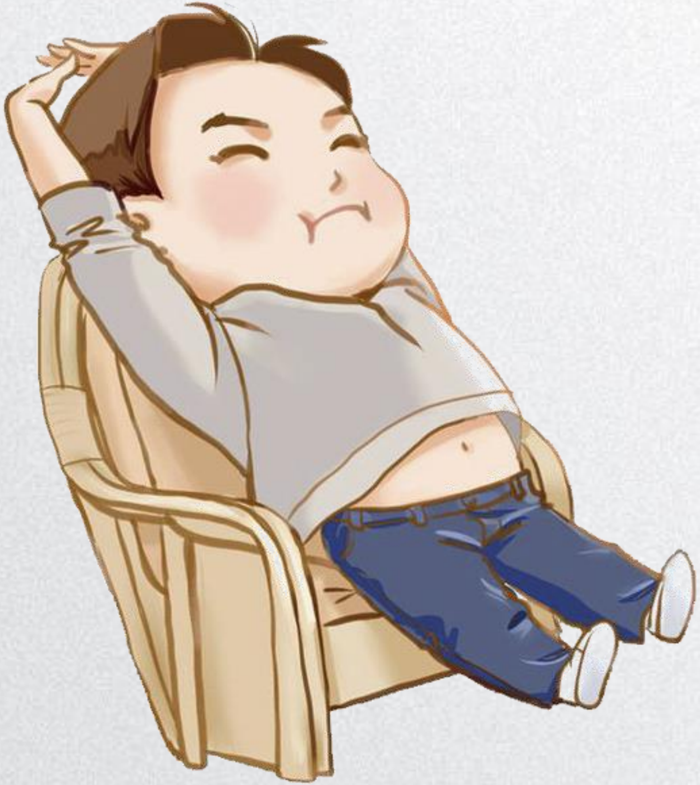
$$\frac{\partial \mathbf{x}^T \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}^T\boldsymbol{x}$$

如果A是对称矩阵

$$2\boldsymbol{A}\boldsymbol{x}$$

π@RUC

矩阵求导



π@RUC