Matlab 14: Matrix Formulas

Cheng-Hsin Hsu

National Tsing Hua University

Department of Computer Science

Slides and sample codes are based on the materials from
Prof. Roger Jang

Transpose and Inverse

- Matrix transpose

$$(AB)^T = B^T A^T$$
$$(ABC)^T = C^T B^T A^T$$
$$\mathbf{a}^T \mathbf{b} = \mathbf{b}^T \mathbf{a}$$
$$\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x}$$

- Matrix inverse

$$(AB)^{-1} = B^{-1} A^{-1}$$
$$(ABC)^{-1} = C^{-1} B^{-1} A^{-1}$$

Block Form of a Matrix (1/2)

- Matrix partition into a block form:
  - Examples

$$A = \left[\begin{array}{c|cc} 1 & 2 & 3 \\ 4 & 5 & 6 \\ \hline 7 & 8 & 9 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$A = \begin{bmatrix} | & & | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_j & \cdots & \mathbf{a}_n \\ | & & | & & | \end{bmatrix} = \begin{bmatrix} - & \mathbf{b}_1^T & - \\ & \vdots & \\ - & \mathbf{b}_i^T & - \\ & \vdots & \\ - & \mathbf{b}_m^T & - \end{bmatrix}$$

Column vector!

Row vector!

Block Form of a Matrix (2/2)

- Block-form matrix operations
  - Examples

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$\Rightarrow A^T = \begin{bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{bmatrix}, AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

$$A = \begin{bmatrix} | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \\ | & | & | \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \Rightarrow A\mathbf{x} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3$$

Gradient of a Function

- Gradient of a function $f(\mathbf{x})$

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \partial f(\mathbf{x}) / \partial x_1 \\ \vdots \\ \partial f(\mathbf{x}) / \partial x_n \end{bmatrix}, \ where\ \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- If $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x} = \mathbf{x}^T \mathbf{c}$

$$\nabla f(\mathbf{x}) = \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Quadratic Form (1)

- Quadratic form of $\mathbf{x}$

$$\mathbf{x}^T A\mathbf{x} = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} a_{ij} x_i x_j$$

- A can be assumed symmetric since

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T \frac{A + A^T}{2} \mathbf{x}$$

Quadratic Form (2)

- When $\mathbf{x} = [x, y]^T$

$$\mathbf{x}^T A \mathbf{x} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x + a_{21}y & a_{12}x + a_{22}y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= a_{11}x^2 + (a_{12} + a_{21})xy + a_{22}y^2$$

- Different values of A can lead to the same quadratic form as long as is the same.

Quadratic Form (3)

- When $\mathbf{x} = [x, y, z]^T$

$$\mathbf{x}^T A \mathbf{x} = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$= a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + (a_{12} + a_{21})xy + (a_{13} + a_{31})xz + (a_{23} + a_{32})yz$$

Expand it by
Brutal force!

- Different values of A can lead to the same quadratic form too.

Gradient of a Quadratic Form

- The gradient of a quadratic form

$$\nabla\left(\mathbf{x}^T A\mathbf{x}\right) = \begin{cases} 2A\mathbf{x}, \textit{ if } A \textit{ is symetric} \\ \left(A + A^T\right)\mathbf{x}, \textit{ otherwise} \end{cases}$$

  – Example

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$= a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + \left(a_{12} + a_{21}\right)xy + \left(a_{13} + a_{31}\right)xz + \left(a_{23} + a_{32}\right)yz$$

$$\nabla\left(\mathbf{x}^T A\mathbf{x}\right) = \begin{bmatrix} 2a_{11}x + \left(a_{12} + a_{21}\right)y + \left(a_{13} + a_{31}\right)z \\ 2a_{22}y + \left(a_{12} + a_{21}\right)x + \left(a_{23} + a_{32}\right)z \\ 2a_{33}z + \left(a_{13} + a_{31}\right)x + \left(a_{23} + a_{32}\right)y \end{bmatrix}$$

$$\Rightarrow \quad = \begin{bmatrix} 2a_{11} & a_{12} + a_{21} & a_{13} + a_{31} \\ a_{12} + a_{21} & 2a_{22} & a_{23} + a_{32} \\ a_{13} + a_{31} & a_{23} + a_{32} & 2a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$= \left(A + A^T\right)\mathbf{x}$$

Common Formulas

- Some common formulas (assuming A is symmetric and all derivatives are w.r.t x)

$$\nabla\left(\mathbf{x}^T A\mathbf{x}\right) = 2A\mathbf{x}$$

$$\nabla\left(\mathbf{x}^T \mathbf{c}\right) = \nabla\left(\mathbf{c}^T \mathbf{x}\right) = \mathbf{c}$$

$$\nabla\left(\mathbf{x}^T \mathbf{x}\right) = 2\mathbf{x}$$

$$\nabla\left(\mathbf{x}^T A\mathbf{c}\right) = A\mathbf{c}$$

$$\nabla\left(\mathbf{c}^T A\mathbf{x}\right) = A^T\mathbf{c}$$

$$\nabla\left(\mathbf{x}^T A\mathbf{x} + \mathbf{b}^T\mathbf{x} + \mathbf{c}\right) = 2A\mathbf{x} + \mathbf{b}$$

Reference

- Matrix cookbook
  - http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf