Review of Matrix Math

CE 130L. Uncertainty, Design, and Optimization
Spring, 2009

Consider column vectors $\mathbf{x}$ and $\mathbf{y}$, both of length n, and a square n by n matrix $\mathbf{Z}$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} , \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} , \quad \text{and} \quad \mathbf{Z} = \begin{bmatrix} Z_{11} & Z_{12} & \dots & Z_{1n} \\ Z_{21} & Z_{22} & \dots & Z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{n1} & Z_{n2} & \dots & Z_{nn} \end{bmatrix}$$

Vector and Matrix Transpose

The *transpose* of the vector $\mathbf{x}$ and the matrix $\mathbf{Z}$ are

$$\mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \quad \text{and} \quad \mathbf{Z}^T = \begin{bmatrix} Z_{11} & Z_{21} & \dots & Z_{n1} \\ Z_{12} & Z_{22} & \dots & Z_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{1n} & Z_{2n} & \dots & Z_{nn} \end{bmatrix} .$$

For column vectors $\mathbf{x}$ and $\mathbf{y}$

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}^T = \begin{bmatrix} \mathbf{x}^T & \mathbf{y}^T \end{bmatrix} .$$

For matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ of appropriate dimensions,

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right]^T = \left[\begin{array}{c|c} \mathbf{A}^T & \mathbf{C}^T \\ \hline \mathbf{B}^T & \mathbf{D}^T \end{array}\right] .$$

Symmetric Matrices

A square matrix $\mathbf{Z}$ in which $Z_{ij} = Z_{ji}$ is called *symmetric*.
For all symmetric matrices $\mathbf{Z} = \mathbf{Z}^T$.

Products

The *scalar product* of the column vectors $\mathbf{x}$ and $\mathbf{y}$ is $\mathbf{x}^T\mathbf{y}$ and is a scalar,

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{y}^T\mathbf{x} \ .$$

The product $\mathbf{Zx}$ of the matrix $\mathbf{Z}$ and the column vector $\mathbf{x}$ is a column vector

$$\begin{aligned}
\mathbf{Zx} &= \begin{bmatrix} Z_{11} & Z_{12} & \ldots & Z_{1n} \\ Z_{21} & Z_{22} & \ldots & Z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{n1} & Z_{n2} & \ldots & Z_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} Z_{11}x_1 + Z_{12}x_2 + \ldots + Z_{1n}x_n \\ Z_{21}x_1 + Z_{22}x_2 + \ldots + Z_{2n}x_n \\ \vdots \\ Z_{n1}x_1 + Z_{n2}x_2 + \ldots + Z_{nn}x_n \end{bmatrix} \\
&= \begin{bmatrix} \sum Z_{1i}x_i \\ \sum Z_{2i}x_i \\ \vdots \\ \sum Z_{ni}x_i \end{bmatrix} \ .
\end{aligned}$$

Likewise, $\mathbf{x}^T\mathbf{Z} = \begin{bmatrix} \sum x_iZ_{i1} & \sum x_iZ_{i2} & \ldots & \sum x_iZ_{in} \end{bmatrix}$

Fact: The transpose of a product is the product of the transposes, $[\mathbf{Zx}]^T = \mathbf{x}^T\mathbf{Z}^T$

Proof:

$$\begin{aligned}
[\mathbf{Zx}]^T &= \begin{bmatrix} \sum Z_{1i}x_i & \sum Z_{2i}x_i & \ldots & \sum Z_{ni}x_i \end{bmatrix} \\
\mathbf{x}^T\mathbf{Z}^T &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} Z_{11} & Z_{21} & \ldots & Z_{n1} \\ Z_{12} & Z_{22} & \ldots & Z_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{1n} & Z_{2n} & \ldots & Z_{nn} \end{bmatrix} \\
&= \begin{bmatrix} x_1Z_{11} + x_2Z_{12} + \cdots + x_nZ_{1n} & x_1Z_{21} + x_2Z_{22} + \cdots + x_nZ_{2n} & \cdots & x_1Z_{n1} + x_2Z_{n2} + \cdots + x_nZ_{nn} \end{bmatrix} \\
&= \begin{bmatrix} \sum x_iZ_{1i} & \sum x_iZ_{2i} & \ldots & \sum x_iZ_{ni} \end{bmatrix} \\
&= [\mathbf{Zx}]^T
\end{aligned}$$

The product $\mathbf{x}^T\mathbf{Z}\mathbf{x}$ is called a *quadratic form* and is a scalar.

$$\mathbf{x}^T\mathbf{Z}\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_k & \cdots & x_n \end{bmatrix} \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1k} & \cdots & Z_{1n} \\ Z_{21} & Z_{22} & \cdots & Z_{2k} & \cdots & Z_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ Z_{k1} & Z_{k2} & \cdots & Z_{kk} & \cdots & Z_{kn} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ Z_{n1} & Z_{n2} & \cdots & Z_{nk} & \cdots & Z_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ \vdots \\ x_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1 & x_2 & \cdots & x_k & \cdots & x_n \end{bmatrix} \begin{bmatrix} Z_{11}x_1 + Z_{12}x_2 + \cdots + Z_{1k}x_k + \cdots + Z_{1n}x_n \\ Z_{21}x_1 + Z_{22}x_2 + \cdots + Z_{2k}x_k + \cdots + Z_{2n}x_n \\ \vdots \\ Z_{k1}x_1 + Z_{k2}x_2 + \cdots + Z_{kk}x_k + \cdots + Z_{kn}x_n \\ \vdots \\ Z_{n1}x_1 + Z_{n2}x_2 + \cdots + Z_{nk}x_k + \cdots + Z_{nn}x_n \end{bmatrix}$$

$$= \begin{array}{l} x_1Z_{11}x_1 + x_1Z_{12}x_2 + \cdots + x_1Z_{1k}x_k + \cdots + x_1Z_{1n}x_n + \\ x_2Z_{21}x_1 + x_2Z_{22}x_2 + \cdots + x_2Z_{2k}x_k + \cdots + x_2Z_{2n}x_n + \cdots + \\ x_kZ_{k1}x_1 + x_kZ_{k2}x_2 + \cdots + x_kZ_{kk}x_k + \cdots + x_kZ_{kn}x_n + \cdots + \\ x_nZ_{n1}x_1 + x_nZ_{n2}x_2 + \cdots + x_nZ_{nk}x_k + \cdots + x_nZ_{nn}x_n \end{array}$$

$$= \begin{bmatrix} x_1 & x_2 & \cdots & x_k & \cdots & x_n \end{bmatrix} \begin{bmatrix} \sum Z_{1i}x_i \\ \sum Z_{2i}x_i \\ \vdots \\ \sum Z_{ki}x_i \\ \vdots \\ \sum Z_{ni}x_i \end{bmatrix}$$

$$= x_1 \sum Z_{1i}x_i + x_2 \sum Z_{2i}x_i + \cdots + x_k \sum Z_{ni}x_i + \cdots + x_n \sum Z_{ni}x_i$$

$$= \sum_{j=1}^{n} x_j \sum_{i=1}^{n} Z_{ji}x_i$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i Z_{ij} x_j$$

Gradients

The derivative of a scalar q with respect to a column vector $\mathbf{x}$ is called the *gradient* and is a row vector.

$$\frac{\partial q}{\partial \mathbf{x}} = \left[\begin{array}{cccc} \frac{\partial q}{\partial x_1} & \frac{\partial q}{\partial x_2} & \cdots & \frac{\partial q}{\partial x_n} \end{array} \right]$$

If

$$q(\mathbf{x}) = \mathbf{y}^T \mathbf{x} = y_1 x_1 + y_2 x_2 + \cdots + y_n x_n = \sum_{i=1}^{n} y_i x_i$$

then

$$\frac{\partial q}{\partial \mathbf{x}} = \left[\begin{array}{cccc} y_1 & y_2 & \cdots & y_n \end{array} \right] = \mathbf{y}^T \ .$$

If

$$\begin{aligned} q(\mathbf{x}) &= \mathbf{x}^T \mathbf{Z} \mathbf{x} \\ &= \begin{array}{l} x_1 Z_{11} x_1 + x_1 Z_{12} x_2 + \cdots + x_1 Z_{1k} x_k + \cdots + x_1 Z_{1n} x_n + \\ x_2 Z_{21} x_1 + x_2 Z_{22} x_2 + \cdots + x_2 Z_{2k} x_k + \cdots + x_2 Z_{2n} x_n + \cdots + \\ x_k Z_{k1} x_1 + x_k Z_{k2} x_2 + \cdots + x_k Z_{kk} x_k + \cdots + x_k Z_{kn} x_n + \cdots + \\ x_n Z_{n1} x_1 + x_n Z_{n2} x_2 + \cdots + x_n Z_{nk} x_k + \cdots + x_n Z_{nn} x_n \end{array} \end{aligned}$$

then

$$\begin{aligned} \frac{\partial q}{\partial x_k} &= x_1 Z_{1k} + x_2 Z_{2k} + \cdots + x_k Z_{kk} + \cdots + x_n Z_{nk} + \\ & \quad Z_{k1} x_1 + Z_{k2} x_2 + \cdots + Z_{kk} x_k + \cdots + Z_{kn} x_n \\ &= \sum_{i=1}^{n} x_i Z_{ik} + \sum_{j=1}^{n} Z_{kj} x_j \end{aligned}$$

and

$$\frac{\partial q}{\partial \mathbf{x}} = \mathbf{x}^T \mathbf{Z} + [\mathbf{Z}\mathbf{x}]^T = \mathbf{x}^T \left[\mathbf{Z} + \mathbf{Z}^T \right] \ .$$

If $q(\mathbf{x}) = \mathbf{x}^T \mathbf{Z} \mathbf{x}$ and $\mathbf{Z}$ is symmetric ($\mathbf{Z} = \mathbf{Z}^T$) then

$$\frac{\partial q}{\partial \mathbf{x}} = 2 \ \mathbf{x}^T \mathbf{Z} \ .$$