Notes on Special Matrices

Paul F. Roysdon, Ph.D.

I. Block matrices

Let $\mathbf{A}_{ij}$ denote the $(i,j)^{th}$ block of the matrix $\mathbf{A}$.

A. Multiplication

Assuming the dimensions of the blocks are equal, we have

$$\left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right]\left[\begin{array}{c|c}\mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22}\end{array}\right] = \left[\begin{array}{c|c}\mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12}+\mathbf{A}_{12}\mathbf{B}_{22} \\ \hline \mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12}+\mathbf{A}_{22}\mathbf{B}_{22}\end{array}\right]$$

B. The Determinant

The determinant can be expressed as

$$\begin{aligned}\det\left(\left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right]\right) &= \det(\mathbf{A}_{22})\cdot\det(\mathbf{C}_1) \\ &= \det(\mathbf{A}_{11})\cdot\det(\mathbf{C}_2)\end{aligned}$$

where,

$$\begin{aligned}\mathbf{C}_1 &= \mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} \\ \mathbf{C}_2 &= \mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\end{aligned}$$

C. The Inverse of a Block Matrix

The inverse can be expressed as

$$\begin{aligned}&\left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right]^{-1} \\ &= \left[\begin{array}{c|c}\mathbf{C}_1^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{C}_2^{-1} \\ \hline -\mathbf{C}_2^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{C}_2^{-1}\end{array}\right] \\ &= \left[\begin{array}{c|c}\mathbf{B}_1 & \mathbf{B}_2 \\ \hline \mathbf{B}_3 & \mathbf{B}_4\end{array}\right]\end{aligned}$$

where,

$$\begin{aligned}\mathbf{B}_1 &= \mathbf{A}_{11}^{-1}+\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{C}_2^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} \\ \mathbf{B}_2 &= -\mathbf{C}_1^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{B}_3 &= -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{C}_1^{-1} \\ \mathbf{B}_4 &= \mathbf{A}_{22}^{-1}+\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{C}_1^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{C}_1 &= \mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} \\ \mathbf{C}_2 &= \mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\end{aligned}$$

D. Block diagonal

For block diagonal matrices we have

$$\left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{A}_{22}\end{array}\right]^{-1} = \left[\begin{array}{c|c}(\mathbf{A}_{11})^{-1} & \mathbf{0} \\ \hline \mathbf{0} & (\mathbf{A}_{22})^{-1}\end{array}\right]$$

$$\det\left(\left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{A}_{22}\end{array}\right]\right) = \det(\mathbf{A}_{11})\cdot\det(\mathbf{A}_{22})$$

E. Schur complement

Consider the matrix

$$\left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right]$$

The Schur complement of block $\mathbf{A}_{11}$ of the matrix above is the matrix (denoted $\mathbf{C}_2$ in the text above)

$$\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}$$

The Schur complement of block $\mathbf{A}_{22}$ of the matrix above is the matrix (denoted $\mathbf{C}_1$ in the text above)

$$\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}$$

Using the Schur complement, one can rewrite the inverse of a block matrix

$$\begin{aligned}\left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right]^{-1} &= \left[\begin{array}{c|c}\mathbf{I} & \mathbf{0} \\ \hline -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I}\end{array}\right]\cdot \\ &\left[\begin{array}{c|c}(\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{A}_{22}^{-1}\end{array}\right]\cdot\left[\begin{array}{c|c}\mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \hline \mathbf{0} & \mathbf{I}\end{array}\right]\end{aligned}$$

The Schur complement is useful when solving linear systems of the form

$$\left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right]\left[\begin{array}{c}\boldsymbol{x}_1 \\ \hline \boldsymbol{x}_2\end{array}\right] = \left[\begin{array}{c}\mathbf{b}_1 \\ \hline \mathbf{b}_2\end{array}\right]$$

which has the following equation for $\boldsymbol{x}_1$

$$(\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})\boldsymbol{x}_1 = \mathbf{b}_1-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{b}_2$$

When the appropriate inverses exists, this can be solved for $\boldsymbol{x}_1$ which can then be inserted in the equation for $\boldsymbol{x}_2$ to solve for $\boldsymbol{x}_2$.

II. Hermitian Matrices and skew-Hermitian

A. Hermitian Matrices

A matrix $\mathbf{A}\in\mathbb{C}^{m\times n}$ is called *Hermitian* if

$$\mathbf{A}^H = \mathbf{A}$$

For real valued matrices, Hermitian and symmetric matrices are equivalent.

$$\begin{aligned}\mathbf{A}\text{ is Hermitian} &\iff \boldsymbol{x}^H\mathbf{A}\boldsymbol{x}\in\mathbb{R},\quad \forall\boldsymbol{x}\in\mathbb{C}^{n\times 1} \\ \mathbf{A}\text{ is Hermitian} &\iff \mathrm{eig}(\mathbf{A})\in\mathbb{R}\end{aligned}$$

Note that

$$\mathbf{A} = \mathbf{B}+i\mathbf{C}$$

where $\mathbf{B}$, $\mathbf{C}$ are Hermitian, then

$$\begin{aligned}\mathbf{B} &= \frac{\mathbf{A}+\mathbf{A}^H}{2} \\ \mathbf{C} &= \frac{\mathbf{A}-\mathbf{A}^H}{2i}\end{aligned}$$

B. *skew-Hermitian*

A matrix $\mathbf{A}$ is called *skew-Hermitian* if

$$\mathbf{A} = -\mathbf{A}^H$$

For real valued matrices, skew-Hermitian and skew-symmetric matrices are equivalent.

$$\begin{aligned}
\mathbf{A} \text{ is Hermitian} &\iff i\mathbf{A} \text{ is skew-Hermitian} \\
\mathbf{A} \text{ is skew-Hermitian} &\iff \boldsymbol{x}^H\mathbf{A}\mathbf{y} = -\boldsymbol{x}^H\mathbf{A}^H\mathbf{y}, \quad \forall \boldsymbol{x}, \mathbf{y} \\
\mathbf{A} \text{ is skew-Hermitian} &\iff \text{eig}(\mathbf{A}) = i\lambda, \quad \lambda \in \mathbb{R}
\end{aligned}$$

III. Idempotent Matrices

A matrix $\mathbf{A}$ is idempotent if

$$\mathbf{A}\mathbf{A} = \mathbf{A}$$

Idempotent matrices $\mathbf{A}$ and $\mathbf{B}$, have the following properties

$$\begin{aligned}
\mathbf{A}^n &= \mathbf{A}, \text{ for } n = 1, 2, 3, \ldots \\
\mathbf{I} - \mathbf{A} &\text{ is idempotent} \\
\mathbf{A}^H &\text{ is idempotent} \\
\mathbf{I} - \mathbf{A}^H &\text{ is idempotent} \\
\text{If } \mathbf{AB} = \mathbf{BA} &\implies \mathbf{AB} \text{ is idempotent} \\
\text{rank}(\mathbf{A}) &= \text{Tr}(\mathbf{A}) \\
\mathbf{A}(\mathbf{I} - \mathbf{A}) &= 0 \\
(\mathbf{I} - \mathbf{A})\mathbf{A} &= 0 \\
\mathbf{A}^+ &= \mathbf{A} \\
f(s\mathbf{I} + t\mathbf{A}) &= (\mathbf{I} - \mathbf{A})f(s) + \mathbf{A}f(s+t)
\end{aligned}$$

Note that $\mathbf{A} - \mathbf{I}$ is not necessarily idempotent.

A. *Nilpotent*

A matrix $\mathbf{A}$ is nilpotent if

$$\mathbf{A}^2 = \mathbf{0}$$

A nilpotent matrix has the following property:

$$f(s\mathbf{I} + t\mathbf{A}) = \mathbf{I}f(s) + t\mathbf{A}f'(s)$$

B. *Unipotent*

A matrix $\mathbf{A}$ is unipotent if

$$\mathbf{A}\mathbf{A} = \mathbf{I}$$

A unipotent matrix has the following property:

$$f(s\mathbf{I} + t\mathbf{A}) = [(\mathbf{I} + \mathbf{A})f(s+t) + (\mathbf{I} - \mathbf{A})f(s-t)]/2$$

IV. Orthogonal Matrices

If a square matrix $\mathbf{Q}$ is orthogonal, if and only if,

$$\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{Q}\mathbf{Q}^\mathsf{T} = \mathbf{I}$$

and then $\mathbf{Q}$ has the following properties

- Its eigenvalues are placed on the unit circle.
- Its eigenvectors are unitary, i.e. have length one.
- The inverse of an orthogonal matrix is orthogonal too.

Basic properties for the orthogonal matrix $\mathbf{Q}$

$$\begin{aligned}
\mathbf{Q}^{-1} &= \mathbf{Q}^\mathsf{T} \\
\mathbf{Q}^{\mathsf{T}} &= \mathbf{Q} \\
\mathbf{Q}\mathbf{Q}^\mathsf{T} &= \mathbf{I} \\
\mathbf{Q}^\mathsf{T}\mathbf{Q} &= \mathbf{I} \\
|\mathbf{Q}| &= \pm 1
\end{aligned}$$

A. *Ortho-Symmetric*

A matrix $\mathbf{Q}_+$ which simultaneously is orthogonal and symmetric is called an ortho-sym matrix. Therefore

$$\begin{aligned}
\mathbf{Q}_+^\mathsf{T}\mathbf{Q}_+ &= \mathbf{I} \\
\mathbf{Q}_+ &= \mathbf{Q}_+^\mathsf{T}
\end{aligned}$$

The powers of an ortho-sym matrix are given by the following rule

$$\begin{aligned}
\mathbf{Q}_+^k &= \frac{1 + (-1)^k}{2}\mathbf{I} + \frac{1 + (-1)^{k+1}}{2}\mathbf{Q}_+ \\
&= \frac{1 + \cos(k\pi)}{2}\mathbf{I} + \frac{1 - \cos(k\pi)}{2}\mathbf{Q}_+
\end{aligned}$$

B. *Ortho-Skew*

A matrix which simultaneously is orthogonal and antisymmetric is called an ortho-skew matrix. Therefore

$$\begin{aligned}
\mathbf{Q}_-^H\mathbf{Q}_- &= \mathbf{I} \\
\mathbf{Q}_- &= \mathbf{Q}_-^H
\end{aligned}$$

The powers of an ortho-skew matrix are given by the following rule

$$\begin{aligned}
\mathbf{Q}_-^k &= \frac{i^k + (-i)^k}{2}\mathbf{I} - i\frac{i^k + (-i)^k}{2}\mathbf{Q}_- \\
&= \cos(k\frac{\pi}{2})\mathbf{I} + \sin(k\frac{\pi}{2})\mathbf{Q}_-
\end{aligned}$$

V. Positive Definite and Semi-definite Matrices

A. *Definition*

A matrix $\mathbf{A}$ is positive definite if and only if

$$\boldsymbol{x}^\mathsf{T}\mathbf{A}\boldsymbol{x} > \mathbf{0}, \ \forall \boldsymbol{x} \neq \mathbf{0}$$

A matrix $\mathbf{A}$ is positive semi-definite if and only if

$$\boldsymbol{x}^\mathsf{T}\mathbf{A}\boldsymbol{x} \geq \mathbf{0}, \ \forall \boldsymbol{x}$$

Note that if $\mathbf{A}$ is positive definite, then $\mathbf{A}$ is also positive semi-definite.

B. *Eigenvalues*

The following holds with respect to the eigenvalues:

$$\begin{aligned}
\mathbf{A} \text{ pos. def.} &\iff \text{eig}\left(\frac{\mathbf{A} + \mathbf{A}^H}{2}\right) > 0 \\
\mathbf{A} \text{ pos. semi-def.} &\iff \text{eig}\left(\frac{\mathbf{A} + \mathbf{A}^H}{2}\right) \geq 0
\end{aligned}$$

C. Trace

The following holds with respect to the trace:

$$\begin{aligned}\mathbf{A} \text{ pos. def.} &\implies \text{Tr}(\mathbf{A}) > 0 \\ \mathbf{A} \text{ pos. semi-def.} &\implies \text{Tr}(\mathbf{A}) \geq 0\end{aligned}$$

D. Inverse

If $\mathbf{A}$ is positive definite, then $\mathbf{A}$ is invertible and $\mathbf{A}^{-1}$ is also positive definite.

E. Diagonal

If $\mathbf{A}$ is positive definite, then $A_{ii} > 0, \ \forall i$.

VI. Symmetric, Skew-symmetric/Antisymmetric

A. Symmetric

The matrix $\mathbf{A}$ is said to be symmetric if

$$\mathbf{A} = \mathbf{A}^\mathsf{T}$$

Symmetric matrices have many important properties, e.g. that their eigenvalues are real and eigenvectors orthogonal.

B. Skew-symmetric/Antisymmetric

The antisymmetric matrix is also known as the skew symmetric matrix. It has the following property from which it is defined

$$\mathbf{A} = -\mathbf{A}^\mathsf{T}$$

Hereby, it can be seen that the antisymmetric matrices always have a zero diagonal. The $n \times n$ antisymmetric matrices also have the following properties.

$$\begin{aligned}|\mathbf{A}^\mathsf{T}| &= |-\mathbf{A}| = (-1)^n|\mathbf{A}| \\ -|\mathbf{A}| &= |-\mathbf{A}| = 0, \text{ if n is odd}\end{aligned}$$

The eigenvalues of an antisymmetric matrix are placed on the imaginary axis and the eigenvectors are unitary.

VII. Single-entry Matrix

The single-entry matrix is very useful when working with derivatives of expressions involving matrices.

A. Definition

The single-entry matrix $\mathbf{J}_{ij} \in \mathbb{R}^{n\times n}$ is defined as the matrix which is zero everywhere except in the entry (i, j) in which it is 1.

B. Swap and Zeros

Assume $\mathbf{A} \in \mathbb{R}^{n\times m}$ and $\mathbf{J}_{ij} \in \mathbb{R}^{m\times p}$

$$\mathbf{A}\mathbf{J}_{ij} = [\mathbf{0} \ \cdots \ \mathbf{0} \ \mathbf{A}_i \ \mathbf{0} \ \cdots \ \mathbf{0}]$$

i.e. a $n\times p$ matrix of zeros with the i-th column of $\mathbf{A}$ in place of the j-th column. Assume $\mathbf{A} \in \mathbb{R}^{n\times m}$ and $\mathbf{J}_{ij} \in \mathbb{R}^{p\times n}$

$$\mathbf{J}_{ij}\mathbf{A} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{A}_j \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}.$$

i.e. an $p \times m$ matrix of zeros with the j-th row of $\mathbf{A}$ in the placed of the i-th row.

VIII. Toeplitz Matrices

A Toeplitz matrix $\mathbf{T}$ is a matrix where the elements of each diagonal is the same. In the $n \times n$ square case, it has the following structure:

$$\begin{aligned}\mathbf{T} &= \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{12} \\ t_{n1} & \cdots & t_{21} & t_{11} \end{bmatrix} \\ &= \begin{bmatrix} t_0 & t_1 & \cdots & t_{n-1} \\ t_{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1 \\ t_{-(n-1)} & \cdots & t_{-1} & t_0 \end{bmatrix}.\end{aligned}$$

A Toeplitz matrix is persymmetric. If a matrix is persymmetric (or orthosymmetric), it means that the matrix is symmetric about its northeast-southwest diagonal (anti-diagonal). Persymmetric matrices are a larger class of matrices, since a persymmetric matrix does not necessarily have a Toeplitz structure.

There are some special cases of Toeplitz matrices. The symmetric Toeplitz matrix is given by:

$$\mathbf{T} = \begin{bmatrix} t_0 & t_1 & \cdots & t_{n-1} \\ t_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1 \\ t_{n-1} & \cdots & t_1 & t_0 \end{bmatrix}.$$

The circular Toeplitz matrix:

$$\mathbf{T} = \begin{bmatrix} t_0 & t_1 & \cdots & t_{n-1} \\ t_{n-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1 \\ t_1 & \cdots & t_{n-1} & t_0 \end{bmatrix}.$$

The upper triangular Toeplitz matrix:

$$\mathbf{T} = \begin{bmatrix} t_0 & t_1 & \cdots & t_{n-1} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1 \\ 0 & \cdots & 0 & t_0 \end{bmatrix},$$

and the lower triangular Toeplitz matrix:

$$\mathbf{T} = \begin{bmatrix} t_0 & 0 & \cdots & 0 \\ t_{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ t_{-(n-1)} & \cdots & t_{-1} & t_0 \end{bmatrix}.$$

The Toeplitz matrix has some computational advantages. The addition of two Toeplitz matrices can be done with $\mathcal{O}(n)$ flops, multiplication of two Toeplitz matrices can be done in $\mathcal{O}(n \ln n)$ flops. Toeplitz equation systems can be solved in

$\mathcal{O}(n^2)$ flops. The inverse of a positive definite Toeplitz matrix can be found in $\mathcal{O}(n^2)$ flops too. The inverse of a Toeplitz matrix is persymmetric. The product of two lower triangular Toeplitz matrices is a Toeplitz matrix.

IX. Vandermonde Matrices

A Vandermonde matrix has the form

$$\mathbf{V} = \begin{bmatrix} 1 & V_1 & v_1^2 & \cdots & v_1^{n-1} \\ 1 & V_2 & v_2^2 & \cdots & v_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & V_n & v_n^2 & \cdots & v_n^{n-1} \end{bmatrix}$$

The transpose of $\mathbf{V}$ is also said to a Vandermonde matrix. The determinant is given by

$$\det \mathbf{V} = \prod_{i>j} (v_i - v_j)$$