A Collection of the Preliminaries and Theorems in Linear Algebra

by: S.-Behzad Nouri

Last Update: Oct. 7, 2009

Table of Contents

Table of Contents Notation		ii	
		iii	
	0.1 Nomenclature	iii	
1	Derivative of Matrices	1	
	1.1 Preliminaries	1	
2	Taylor's Expansion for Matrix Function	2	
	2.1 Preliminaries	2	
3	Eigen-Values and Diagonal Factorization	5	
4	The Singular Value decomposition (SVD)	8	
5	Bounded Real Lemma (BRL)	10	
	5.1 Preliminaries	10	
6	Diagonally Dominant Matrices	12	
Lis	List of References		

Notation

0.1 Nomenclature

 \mathbb{R} The field of real numbers

 \mathbb{R}^N The set of real vectors of size N

 $\mathbb{R}^{N\times N}$ The set of real matrices of size N × N

 \mathbb{C} The field of complex numbers, e.g.: s-plane

 \mathbb{C}^+ The open right half plane in the complex s-plane; $\mathbb{C}^+ = \{s \in \mathbb{C} : \operatorname{Re}(s) > 0\}$

 \mathbb{C}^- The open left half plane in the complex s-plane; $\mathbb{C}^-=\{s\in\mathbb{C}:\,\mathrm{Re}(s)\ < 0\}$

 $\overline{\mathbb{C}}$ The closed right half plane in the complex s-plane; $\overline{\mathbb{C}}=\{s\in\mathbb{C}:\ \mathrm{Re}(s)\geq 0\}$

 \mathbb{F} The field of complex or real numbers

 $spec(\mathbf{A})$ spectrum of \mathbf{A} , the set of eigenvalues of the matrix $\{\lambda_i\}$

 $\boldsymbol{mspec}\left(\cdot\right)$ is read as $\{\lambda_i\}$, even if it is multiple

 $\boldsymbol{tr}(\cdot)$ the trace, i.e., the sum of the diagonal elements of its argument

 \mathbf{A}^* for $\mathbf{A} \in \mathbb{F}^{n \times n}$ the conjugate (Hermitian) transpose of $\mathbf{A} = [a_{ij}]$ as $a_{ij}^* = (a_{ji})^*$, where $\mathbf{A}^* = [a_{ij}^*]$

Derivative of Matrices

1.1 Preliminaries

Theorem 1.1.1 Let $\mathbf{A} = [a_{ij}]$ be a square nonsingular (hence, invertible) matrix of order n. If its inverse \mathbf{A}^{-1} is derivable with respect to one of its parameters (e.g. x), it is

$$\frac{d}{dx}\left(\mathbf{A}^{-1}\right) = -\mathbf{A}^{-1}\left(\frac{d}{dx}A\right)\mathbf{A}^{-1} \tag{1.1}$$

Proof: Consider $AA^{-1} = I$,

$$\frac{d}{dx} \left(\mathbf{A} \mathbf{A}^{-1} \right) = \left(\frac{d}{dx} \mathbf{A} \right) \mathbf{A}^{-1} + \mathbf{A} \left(\frac{d}{dx} \mathbf{A}^{-1} \right) = \mathbf{0}$$

$$\mathbf{A} \left(\frac{d}{dx} \mathbf{A}^{-1} \right) = -\left(\frac{d}{dx} \mathbf{A} \right) \mathbf{A}^{-1}$$

$$\frac{d}{dx} \left(\mathbf{A}^{-1} \right) = -\mathbf{A}^{-1} \left(\frac{d}{dx} \mathbf{A} \right) \mathbf{A}^{-1}$$

Note: As an special case, for **A** as a scalar variable it is $\frac{d}{dx} \left(\frac{1}{\mathbf{A}} \right) = -\frac{A'}{A^2}$, which can be written in a form as $\left(\frac{1}{\mathbf{A}} \right)' = -\mathbf{A}^{-1}\mathbf{A}'\mathbf{A}^{-1}$.

Taylor's Expansion for Matrix Function

2.1 Preliminaries

Remark 2.1.1 Let $F(s) = (\mathbf{I} - s\mathbf{A}^{-1})^{-1}$ be a smooth (continuously derivable, C^{∞}) matrix function, its Taylor expansion in proximity of s = 0 is

$$F(s) = \sum_{i=0}^{\infty} \mathbf{M}_i s^i = \sum_{i=0}^{\infty} \frac{1}{i!} F^{(i)}(s)|_{s=0} s^i, \qquad (2.1)$$

where $\frac{1}{i!}F^{(i)}(s)|_{s=0} = \mathbf{A}^{-i}$

Proof: consider the subsequent derivatives as

i.
$$F(s) = (\mathbf{I} - s\mathbf{A}^{-1})^{-1} \longrightarrow F(0) = F(s)|_{s=0} = \mathbf{I},$$

ii .
$$F^{(1)}(s) = -F(s) \frac{d}{ds} (\mathbf{I} - s\mathbf{A}^{-1}) F(s) = F(s) \mathbf{A}^{-1} F(s) \longrightarrow F^{(1)}(s)|_{s=0} = F(0) \mathbf{A}^{-1} F(0) = \mathbf{A}^{-1}$$

iii .
$$F^{(2)}(s) = F^{(1)}(s)\mathbf{A}^{-1}F(s) + F(s)\mathbf{A}^{-1}F^{(1)}(s) \longrightarrow F^{(2)}(s)|_{s=0} = F^{(1)}(0)\mathbf{A}^{-1}F(0) + F(0)\mathbf{A}^{-1}F^{(1)}(0) = 2\mathbf{A}^{-2}$$

$$\mathbf{iv} \cdot F^{(3)}(s) = F^{(2)}(s)\mathbf{A}^{-1}F(s) + F^{(1)}(s)\mathbf{A}^{-1}F^{(1)}(s) + F^{(1)}(s)\mathbf{A}^{-1}F^{(1)}(s) + F(s)\mathbf{A}^{-1}F^{(2)}(s) \longrightarrow F^{(3)}(s)|_{s=0} = 2\mathbf{A}^{-3} + \mathbf{A}^{-3} + \mathbf{A}^{-3} + 2\mathbf{A}^{-3} = 6\mathbf{A}^{-3}$$

 \mathbf{v} . and so on so force!

Plunging all above evaluated derivatives in (2.1) results in (2.2) below.

$$F(s) = (\mathbf{I} - s\mathbf{A}^{-1})^{-1} \approx \mathbf{I} + \mathbf{A}^{-1}s + \mathbf{A}^{-2}s^{2} + \mathbf{A}^{-3}s^{3} + \dots$$
 (2.2)

Theorem 2.1.2 Let p be a polynomial and suppose A and B are squared matrices of the same size, then $p(\mathbf{A} + \mathbf{B}) = \sum_{k=0}^{n} \frac{1}{k!} p^{(k)}(\mathbf{A}) \mathbf{B}^{k}$, where $n = \deg(p)$.

Proof: Since p is a polynomial, we can apply the Taylor expansion:

$$p(x) = \sum_{k=0}^{n} \frac{1}{k!} p^{(k)}(x_0) (x - x_0)^k ,$$

where $n = \deg(p)$. Now let $\mathbf{X} = \mathbf{A} + \mathbf{B}$ and $\mathbf{X}_0 = \mathbf{A}$. The Taylor expansion can be checked as follows: let $p(x) = \sum_{k=0}^{n} a_k x^k$ for coefficients a_k (note that this coefficients can be taken from the space of square matrices defined over a field). We define the formal derivative of this polynomial as $p^{(1)}(x) = \frac{dp}{dx} = \sum_{k=1}^{n} a_k k x^{k-1}$ and we define $p^{(k)} = \frac{dp^{(k-1)}}{dx} \frac{1}{k!} p^{(k)}(x_0) = \sum_{i=k}^{n} a_i \frac{i!}{(i-k)!k!} (x_0)^{i-k}$.

Then $p^{(k)}(x) = \sum_{i=k}^{n} a_i \frac{i!}{(i-k)!} x^{i-k}$ and we have $\frac{1}{k!} p^{(k)}(x_0) = \sum_{i=k}^{n} a_i \frac{i!}{(i-k)!k!} (x_0)^{i-k}$. Now consider

$$\sum_{k=0}^{n} \frac{1}{k!} p^{(k)}(x_0) (x - x_0)^k = \sum_{k=0}^{n} \left(\sum_{i=k}^{n} a_i \frac{i!}{(i-k)!k!} (x_0)^{i-k} (x - x_0)^k \right)$$

$$= \sum_{i=0}^{n} a_{i} (x_{0})^{i} + \sum_{i=1}^{n} a_{i} i (x_{0})^{i-1} (x - x_{0}) + \dots + \sum_{i=j}^{n} a_{i} \frac{i!}{(i-j)! j!} (x_{0})^{i-j} (x - x_{0})^{j} + \dots + a_{n} (x - x_{0})^{n}$$

$$= a_{0} + a_{1} (x) + \dots + a_{i} \left(\sum_{i=0}^{i} \frac{i!}{(i-j)! j!} (x_{0})^{i-j} (x - x_{0})^{j} \right) + \dots +$$

$$a_n \left(\sum_{j=0}^n \frac{n!}{(n-j)!j!} (x_0)^{n-j} (x - x_0)^j \right) = \sum_{k=0}^n a_k x^i = p(x)$$

since
$$\sum_{j=0}^{i} \frac{i!}{(i-j)!j!} (x_0)^{i-j} (x-x_0)^j = (x)^i$$
.

Note: To prove theorem 2.1.2 in above [1] has been consulted with.

Eigen-Values and Diagonal Factorization

Definition 3.0.3 For Square matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ define the following types:

- (i) A is **Hermitian** if $A = A^*$.
- (ii) A is positive-semidefinite $(A \ge 0)$ if A is Hermitian and $\mathbf{x}^* A \mathbf{x} \ge 0$ for all $\mathbf{x} \in \mathbb{F}^n$.
- [2], Definition 3.1.1, pp. 81-82

Fact 3.0.4 Consider $\mathbf{A} \in \mathbb{F}^{n \times n}$ (e.g. a scattering parameter matrix of a n-port linear system), $\mathbf{A}^*\mathbf{A}$ and $\mathbf{A}\mathbf{A}^*$ are Hermitian.

Proof:

$$(\mathbf{A}^*\mathbf{A})^* = \mathbf{A}^* (\mathbf{A}^*)^* = \mathbf{A}^* \mathbf{A}$$

For AA^* matrix it is proved similarly.

Fact 3.0.5 Consider $A \in \mathbb{F}^{n \times n}$, matrix A^*A is positive-definite.

Proof: let $\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{F}^n$, then

$$\mathbf{x}^* \left(\mathbf{A}^* \mathbf{A} \right) \mathbf{x} \; = \; \left(\mathbf{x}^* \mathbf{A}^* \right) \left(\mathbf{A} \mathbf{x} \right) \; = \; \left(\mathbf{A} \mathbf{x} \right)^* \left(\mathbf{A} \mathbf{x} \right)$$

It is noted that $(\mathbf{A}\mathbf{x})$ is a column vector with complex entries $(\in \mathbb{F}^n)$. Therefore, we have

$$(\mathbf{A}\mathbf{x})^* (\mathbf{A}\mathbf{x}) =$$

$$\begin{bmatrix} A_{x_1}^* & \dots & A_{x_n}^* \end{bmatrix} \begin{bmatrix} a_{x_1} \\ \vdots \\ a_{x_n} \end{bmatrix} = |a_{x_1}|^2 + \dots + |a_{x_n}|^2 \ge 0.$$

Proposition 3.0.6 Let $\mathbf{A} \in \mathbb{F}^{n \times n}$ and $\alpha \in \mathbb{F}$, then, the following statements hold:

- (i) $mspec(\alpha \mathbf{A}) = \alpha mspec(\mathbf{A})$.
- (ii) $mspec(\beta I_n + \alpha \mathbf{A}) = \beta + \alpha mspec(\mathbf{A}).$
- (iii) if **A** is Hermitian, spec (**A**) $\subset \mathbb{R}$.
- [2], Proposition 4.4.4, pp. 131

In a general form, let $\mathbf{F} \in \mathbb{F}^{n \times m}$, Noting (3.0.5) and (3.0.6), it is concluded

Fact 3.0.7 Matrices $\mathbf{A}^*\mathbf{A} \in \mathbb{F}^{m \times m}$ and $\mathbf{A}\mathbf{A}^* \in \mathbb{F}^{n \times n}$ have **positive-real** eigenvalues.

$$spec\left(\mathbf{A}^{*}\mathbf{A}\right)\subset\mathbb{R}^{+}$$

Fact 3.0.8 Let $\mathbf{A} \in \mathbb{F}^{(n \times n)}$,

$$\|\mathbf{A}\|_F^2 = tr(\mathbf{A}\mathbf{A}^*) = tr(\mathbf{A}^*\mathbf{A}) = \sum_{i,j} |a_{ij}|^2$$

Proof: e.g.:

$$\mathbf{A}^* \mathbf{A} = \begin{bmatrix} a_{11}^* & a_{21}^* \\ a_{12}^* & a_{22}^* \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} |a_{11}^2| + |a_{21}^2| & X \\ X & |a_{12}^2| + |a_{22}^2| \end{bmatrix}$$

$$tr(\mathbf{A}^* \mathbf{A}) = |a_{11}^2| + |a_{21}^2| + |a_{12}^2| + |a_{22}^2| = \|\mathbf{A}\|_F^2.$$

Fact 3.0.9 For the norm of the matrix A we have

$$(i) \ \left\|A\right\|_2 = \sqrt{\max(eig(A^*A))}.$$

$$(ii) \|A\|_{\infty} = \max_{i} \sum_{j} |A_{ij}|.$$

[3], ch. 10.3

Fact 3.0.10 For the (ii) in fact. 3.0.9 and def. 4.0.12, it is norm of the matrix A we have

$$||A||_2 = \sigma_{max}(A). \tag{3.1}$$

The Singular Value decomposition (SVD)

In this section, we briefly review the singular value decomposition method.

Definition 4.0.11 For $\mathbf{A} \in \mathbb{F}^{(n \times m)}$, the SVD decomposition of A is

$$\mathbf{A} = \mathbf{U}_{(n \times n)} \Sigma_{(n \times m)} \mathbf{V}_{(m \times m)}^{T}$$

or equivalently,

$$\mathbf{A} = \mathbf{U}_{(n \times n)} \begin{bmatrix} \Sigma_{(r \times r)} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{V}_{(m \times m)}^{T}$$

where $r = min\{n, m\}$ for a full (column and row) rank \mathbf{A} , $\mathbf{U}_{(n \times n)}$ and $\mathbf{V}_{(\mathbf{m} \times \mathbf{m})}$ are orthogonal matrices, $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$. $\Sigma = Diag\{\sigma_1, \sigma_1, \cdots, \sigma_r\}$, σ_i is called singular values and are defined in c.f. 4.0.12.

When the matrix is not (e.g.) full column rank $r \leq m$, where $r = rank(\mathbf{A})$, a short-en (economy) form of SVD can be

$$\mathbf{A} = \left[\begin{array}{cccc} u_1 & \cdots & u_i & \cdots & u_r \end{array} \right]_{(n \times r)} \Sigma_{(r \times r)} \left[\begin{array}{cccc} v_1 & \cdots & v_i & \cdots & v_r \end{array} \right]_{(r \times m)}^T.$$

Definition 4.0.12 Let $\mathbf{A} \in \mathbb{F}^{(n \times m)}$. Then, the **singular values** of \mathbf{A} are the $min\{m,n\}$ nonnegative number $\sigma_1(A), \cdots, \sigma_{min\{m,n\}}(A)$, where, for all $i=1,\cdots, min\{m,n\}$,

$$\sigma_i(\mathbf{A}) \triangleq \left[\lambda \left(\mathbf{A}\mathbf{A}^*\right)\right]^{1/2} = \left[\lambda \left(\mathbf{A}^*\mathbf{A}\right)\right]^{1/2}.$$

Then,

$$\sigma_1(\mathbf{A}) \geq \cdots \geq \sigma_{\min\{n,m\}}(\mathbf{A}) \geq 0.$$

[2], Definition 5.6.1, pp. 181-182 SVD can lead to the best approximation in terms of the 2-norm.

Fact 4.0.13 Let $\mathbf{A} \in \mathbb{F}^{(n \times m)}$, and let $r = rank\mathbf{A}$. Then, for all $i = 1, \dots, r$,

$$\sigma_i(\mathbf{A}^*\mathbf{A}) = \sigma_i(\mathbf{A}\mathbf{A}^*) = \sigma_i^2(\mathbf{A}).$$

In particular,

$$\sigma_1(\mathbf{A}^*\mathbf{A}) = \sigma_{max}^2(\mathbf{A}).$$

[2], Fact 5.10.18, pp. 198

Fact 4.0.14 Let $\mathbf{A} \in \mathbb{F}^{(n \times n)}$,

$$\|\mathbf{A}\|_F^2 = tr(\mathbf{A}\mathbf{A}^*) = tr(\mathbf{A}^*\mathbf{A}) = \sum_{i,j} |a_{ij}|^2$$

Fact 4.0.15 For the (ii) in fact. 3.0.9 and def. 4.0.12, it is norm of the matrix A we have

$$||A||_2 = \sigma_{max}(A). \tag{4.1}$$

Bounded Real Lemma (BRL)

5.1 Preliminaries

As a (passivity) checking criterion for bounded-real-ness of s-parameter matrix $\mathbf{S} \in \mathbb{F}^{n \times n}$, which is the scattering-parameters for a n-port linear system, we have:

$$I_n - \mathbf{S}^* \mathbf{S} \ge 0. \tag{5.1}$$

From the mathematical elaborations in previous chapter, it is concluded that, eq. (5.1) requires

$$mspec(I - \mathbf{S}^*\mathbf{S}) = 1 - mspec(\mathbf{S}^*\mathbf{S}) \ge 0,$$
 (5.2)

which is equivalently shortened as

$$mspec\left(\mathbf{S}^{*}\mathbf{S}\right) \le 1. \tag{5.3}$$

The above notation in (5.3) shows that all eigenvalues of Hermitian matrix S^*S requires to be bounded to one, while they are known from previous section as real and

positive values. This is logically means that

$$\lambda_{max}\left(\mathbf{S}^*\mathbf{S}\right) \le 1. \tag{5.4}$$

According to the Definition. 4.0.12 we have $[\lambda(\mathbf{A}^*\mathbf{A})] = \sigma_i^2(\mathbf{A})$, by substituting which in (5.3) it is

$$\sigma_{max}^2(\mathbf{S}) \le 1, \tag{5.5}$$

or equivalently

$$\sigma_{max}\left(\mathbf{S}\right) \le 1. \tag{5.6}$$

Fact 5.1.1 Considering fact 4.0.15 and (5.6) the following inequalities are equivalent.

- (i) $\sigma_{max}(\mathbf{S}) \leq 1$.
- (ii) $\|\mathbf{S}\|_2 \leq 1$.

Diagonally Dominant Matrices

Definition 6.0.2 A diagonally dominant matrix is a complex matrix $\Phi = [\varphi_{ij}] \in \mathbb{Z}^{n \times n}$ with the property that we have

$$|\varphi_{ii}| \geqslant \sum_{i=1, i \neq j}^{n} |\varphi_{ij}|, \qquad (6.1)$$

for all i. When all these inequalities are strict, the matrix is called strictly diagonally dominant [4]. where $\| \|$ defines the matrix norm subordinate to the ordinary Euclidean norm for vectors.

Lemma 6.0.3 A symmetric diagonally dominant real matrix with nonnegative diagonal entries is positive semidefinite.

Proof: Given any eigenvalue of matrix λ and corresponding eigenvector \mathbf{U} ,

$$\Phi \mathbf{U} = \lambda \mathbf{U} \tag{6.2}$$

More precisely, it is

and

$$(\varphi_{ii} - \lambda) u_{\mathbf{i}} = -\sum_{i \neq j} \varphi_{ij} u_j.$$
 (6.4)

From (6.4),

$$|(\varphi_{ii} - \lambda) u_{\mathbf{i}}| = \left| \sum_{i \neq j} \varphi_{ij} u_{j} \right|$$

$$\left| \sum_{i \neq j} \varphi_{ij} u_{j} \right| \leqslant \sum_{i \neq j} |\varphi_{ij} u_{j}|$$

$$\sum_{i \neq j} |\varphi_{ij} u_{j}| = \sum_{i \neq j} |\varphi_{ij}| |u_{j}|,$$

$$(6.5)$$

then

$$|(\varphi_{ii} - \lambda) u_{\mathbf{i}}| \leqslant |u_j| \sum_{i \neq j} |\varphi_{ij}|. \tag{6.6}$$

let $|u_{\mathbf{i}}| > |u_{j}|$ for all $j \neq i$, it is

$$|u_j| \sum_{i \neq j} |\varphi_{ij}| < |u_i| \sum_{i \neq j} |\varphi_{ij}| \tag{6.7}$$

(6.6) and (6.7) lead to

$$|(\varphi_{ii} - \lambda) u_{\mathbf{i}}| \leqslant |u_i| \sum_{i \neq j} |\varphi_{ij}|. \tag{6.8}$$

It is equivalently

$$|\varphi_{ii} - \lambda| |u_{\mathbf{i}}| \le |u_i| \sum_{i \ne j} |\varphi_{ij}|,$$
 (6.9)

considering $|u_i| > 0$,

$$|\varphi_{ii} - \lambda| \leqslant \sum_{i \neq j} |\varphi_{ij}| . \tag{6.10}$$

Let assume there exist a negative eigenvalue $\lambda < 0$, then intuitively it is

$$|\varphi_{ii}| < |\varphi_{ii} - \lambda| \tag{6.11}$$

From (6.10) and (6.11), it is

$$|\varphi_{ii}| < \sum_{i \neq j} |\varphi_{ij}| . \tag{6.12}$$

Eq (6.12) is in contradiction with the presumption of diagonally dominance in (6.1), than the assumption of $\lambda < 0$ is not true and eigenvalues are confined to the non-negative values. \Box

Lemma 6.0.4 A strictly diagonally dominant complex matrix is nonsingular.

Lemma 6.0.5 The multiplicity of the eigenvalue 0 of a symmetric real matrix Φ with zero row sums and nonpositive off-diagonal entries equals the number of connected components of the graph Γ defined on the index set of the rows and columns of Φ , where two distinct indices i and j are adjacent when $\varphi_{ij} \neq 0$.

List of References

- [1] The Matrix Cookbook. Available online:

 http://planetmath.org/encyclopedia/ProofOfTaylorsFormulaForMatrixFunctions.htm,
 Sept., 2009.
- [2] D. S. Bernstein, Matrix Mathematics: Theorey, Facts, and Formulas with Application to Linear System Theory. NJ:Priceton University Press, 2005.
- [3] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. Available online: http://matrixcookbook.com, Nov. 14, 2008.
- [4] w, Diagonally dominant matrices. Available online: http://www.win.tue.nl/aeb/srgbk/node16.html: Mathematics Department, Macquarie University, Sydney, Australia, 2002.