MATH36001 Background Material 2015

1 Matrix Algebra

1.1 Matrices and Vectors

An ordered array of mn elements a_{ij} $(i = 1, \dots, m;\ j = 1, \dots, n)$ written in the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

is said to be an $m \times n$ **matrix**. These elements can be taken from an arbitrary field $\mathbb{F}$. However for the purpose of this course, $\mathbb{F}$ will always be the set of all real or all complex numbers, denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. An $m \times n$ matrix may also be written in terms of its elements as

$$A = [a_{ij}],$$

where a_{ij} $(1 \le i \le m,\ 1 \le j \le n)$ denotes the element in position (i, j). Note that i is the row number, and j is the column number. $\mathbb{R}^{m \times n}$ denotes the vector space of all real $m \times n$ matrices, and $\mathbb{C}^{m \times n}$ is the vector space of all complex $m \times n$ matrices. If $m = n$ the matrix is **square**, and in the general case it is **rectangular**.

A **submatrix** of A is any matrix obtained by deleting rows and columns. A **block matrix**

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pq} \end{bmatrix}$$

is a partitioning of A into submatrices A_{ij} whose dimensions must be consistent.

A **row vector** $[\, x_1 \quad x_2 \quad \cdots \quad x_n \,]$ is a $1 \times n$ matrix. A **column vector** $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$ is an $m \times 1$ matrix. Note that $\mathbb{F}^m \equiv \mathbb{F}^{m \times 1}$. $\mathbb{R}^m$ and $\mathbb{C}^m$ denote the vector space of real and complex column vectors with m elements, respectively.

1.2 Zero and Identity Matrices

The **zero matrix** $O_{m \times n}$ is the $m \times n$ matrix all of whose elements are zero. When m and n are clear from the context we may simply write O.

The $n \times n$ **identity matrix** is

$$I_n = \begin{bmatrix} 1 & \dots & \dots & 0 \\ \vdots & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & 1 \end{bmatrix},$$

or just I if the dimension is clear from the context. The jth column of I is denoted by e_j, and is called the jth **unit vector**, that is,

$$I = [\, e_1 \quad e_2 \quad \cdots \quad e_n \,], \qquad e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \; \ldots, \; e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

1.3 Householder's Notation

Generally, we use

capital letters	A, B, C, Δ, Λ	for matrices,
lower case letters	$a_{ij}, b_{ij}, c_{ij}, \delta_{ij}, \lambda_{ij}$	for matrix elements,
lower case letters	x, y, z, c, g, h	for vectors,
lower case Greek letters	$\alpha, \beta, \gamma, \theta, \pi$	for scalars.

1.4 Basic Manipulations with Matrices

Transposition: $(\mathbb{F}^{m\times n} \longrightarrow \mathbb{F}^{n\times m})$

$$C = A^T \iff c_{ij} = a_{ji}.$$

A^T has rows and columns interchanged, so it is an $n \times m$ matrix.
Conjugate transposition: $(\mathbb{F}^{m\times n} \longrightarrow \mathbb{F}^{n\times m})$

$$C = A^* \iff c_{ij} = \overline{a_{ji}},$$

where the bar denotes complex conjugate. If $A \in \mathbb{R}^{m\times n}$ then $A^* = A^T$.
Properties of transposition:

$$\begin{array}{ll} (A^T)^T = A, & (A^*)^* = A, \\ (\alpha A)^T = \alpha A^T, & (\alpha A)^* = \overline{\alpha} A^*, \\ (A+B)^T = A^T + B^T, & (A+B)^* = A^* + B^*, \\ (AB)^T = B^T A^T, & (AB)^* = B^* A^*. \end{array}$$

Addition: $(\mathbb{F}^{m\times n} \times \mathbb{F}^{m\times n} \longrightarrow \mathbb{F}^{m\times n})$

$$C = A + B \iff c_{ij} = a_{ij} + b_{ij}.$$

Scalar-matrix multiplication: $(\mathbb{F} \times \mathbb{F}^{m\times n} \longrightarrow \mathbb{F}^{m\times n})$

$$C = \alpha A \iff c_{ij} = \alpha a_{ij}.$$

Properties of matrix addition:

$A + B = B + A$	matrix addition is commutative
$(A + B) + C = A + (B + C)$	matrix addition is associative
$\alpha(A + B) = \alpha A + \alpha B, \quad (\alpha + \beta)A = \alpha A + \beta A$	matrix addition and scalar multiplication are distributive

Matrix-matrix multiplication: $(\mathbb{F}^{m\times r} \times \mathbb{F}^{r\times n} \longrightarrow \mathbb{F}^{m\times n})$

$$C = AB \iff c_{ij} = \sum_{k=1}^{r} a_{ik}b_{kj}.$$

Properties of matrix multiplication:

$A(BC) = (AB)C$	matrix multiplication is associative
$A(B+C) = AB + AC$	matrix multiplication is distributive

Remember that $AB \neq BA$ in general, that is, matrix multiplication is not commutative.

Matrix powers:
If A is a nonzero square matrix we define $A^0 \equiv I$, and for any positive integer,

$$A^k = \overbrace{A\cdots A}^{k \text{ times}} = A^{k-1}A = AA^{k-1}.$$

A square matrix is **involutory** if $A^2 = I$, **idempotent** if $A^2 = A$, and **nilpotent** if $A^k = 0$ for some integer $k > 0$.

Using powers of matrices, we can also define polynomials in matrices: if $p(z) = c_0 + c_1z + \cdots + c_kz^k$, then given $A \in \mathbb{C}^{n\times n}$, we define $p(A) \in \mathbb{C}^{n\times n}$ by

$$p(A) = c_0I + c_1A + \cdots + c_kA^k.$$

1.5 Inner and Outer Products

The **inner product** of two vectors $x, y \in \mathbb{C}^n$ is the scalar

$$x^*y = \sum_{i=1}^{n} \overline{x_i}y_i \in \mathbb{C}.$$

The **length** of a vector x is given by $\sqrt{x^*x}$.

Two nonzero vectors are **orthogonal** if their inner product x^*y is zero. If, in addition, $x^*x = y^*y = 1$, the vectors are **orthonormal**.

The **outer product** of the vectors $x \in \mathbb{C}^m$ and $y \in \mathbb{C}^n$ is the $m \times n$ matrix

$$xy^* = \begin{bmatrix} x_1\overline{y_1} & \dots & x_1\overline{y_n} \\ \vdots & & \vdots \\ x_m\overline{y_1} & \dots & x_m\overline{y_n} \end{bmatrix} \in \mathbb{C}^{m\times n}.$$

1.6 Special Matrices

A **diagonal matrix** D has all its off-diagonal elements zero. It can be written $D = \mathrm{diag}(\alpha_i)$, or equivalently

$$D = \begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{bmatrix}.$$

Here and below, a blank entry is understood to be zero.

An **upper triangular matrix** U has zero elements below the diagonal, that is $u_{ij} = 0$ for $i > j$, $U = \begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{bmatrix}$, where $\times$ is not necessarily zero.

Similarly, the matrix L is **lower triangular** if all elements above the main diagonal are zero. The matrix A is **block diagonal**, **block upper triangular**, **block lower triangular** if it has the partitioned forms

$$A = \begin{bmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{pp} \end{bmatrix}, \quad A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1p} \\ & A_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & A_{pp} \end{bmatrix}, \quad A = \begin{bmatrix} A_{11} & & & \\ A_{21} & A_{22} & & \\ \vdots & & \ddots & \\ A_{p1} & \cdots & \cdots & A_{pp} \end{bmatrix}.$$

Here the matrices A_{ii} are all square but do not necessarily have the same size.

$A \in \mathbb{R}^{n\times n}$ is a **symmetric matrix** if $A^T = A$; $A \in \mathbb{C}^{n\times n}$ is a **Hermitian matrix** if $A^* = A$. An $n \times n$ Hermitian matrix A is

$$\begin{array}{rll} \text{positive definite} & \text{if} & x^*Ax > 0 \text{ for all } 0 \neq x \in \mathbb{C}^n \\ \text{positive semi-definite} & \text{if} & x^*Ax \geq 0 \text{ for all } x \in \mathbb{C}^n \\ \text{indefinite} & \text{if} & (x^*Ax)(y^*Ay) < 0 \text{ for some } x, y \in \mathbb{C}^n \end{array}$$

An **orthogonal matrix** $Q \in \mathbb{R}^{n\times n}$ satisfies $QQ^T = I$ and $Q^TQ = I$, so that if $Q = [q_1, \ldots, q_n]$ then

$$q_i^T q_j = \delta_{ij} \quad \text{(Kronecker delta)},$$

where $\delta_{ij} = 1$ if $i = j$, 0 otherwise. The columns of Q are mutually orthogonal and of unit length. A **unitary matrix** $U \in \mathbb{C}^{n\times n}$ satisfies $UU^* = U^*U = I$.

The **permutation matrix** P_{ij} is the identity matrix with its ith and jth rows interchanged.

$$P_{ij} = \begin{bmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & 0 & & & 1 & & & \\ & & & & \ddots & & & & & \\ & & & & & 1 & & & & \\ & & & & & & \ddots & & & \\ & & & 1 & & & 0 & & & \\ & & & & & & & 1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 1 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i\text{th row} \\ \\ \\ \\ \leftarrow j\text{th row} \\ \\ \\ \\ \end{array}$$

$P_{ij}A$ swaps the ith and jth rows of A. AP_{ij} swaps the ith and jth columns of A. Note that P_{ij} is orthogonal and involutory ($P_{ij}^2 = I$).

2 Basic Linear Algebra Definitions

The **rank** of a matrix A, $\mathrm{rank}(A)$, is the maximum number of linearly independent rows or columns of A. Recall that a set of vectors $\{v_i\}$ is **linearly dependent** if $\sum_i \alpha_i v_i = 0$ for some scalar values α_i not all zero, and otherwise **linearly independent**.

Two important subspaces associated with a matrix $A \in \mathbb{C}^{m\times n}$ are

$$\begin{aligned} \mathrm{range}(A) &= \big\{y \in \mathbb{C}^m : y = Ax \text{ for some } x \in \mathbb{C}^n\big\}, \quad \text{the } \textbf{range} \text{ of } A \\ \mathrm{null}(A) &= \big\{x \in \mathbb{C}^n : Ax = 0\big\}, \quad \text{the } \textbf{null space} \text{ of } A. \end{aligned}$$

If $A = [a_1, a_2, \dots, a_n]$ (i.e., a_j is the jth column of A), then

$$\begin{aligned} \mathrm{range}(A) &= \mathrm{span}\{a_1, a_2, \dots, a_n\}, \\ \mathrm{rank}(A) &= \dim(\mathrm{range}(A)), \end{aligned}$$

where $\mathrm{span}(S)$ denotes the set of all linear combinations of vectors in the set S, and $\dim(V)$ is the maximum number of linearly independent vectors in the vector space V. For any $A \in \mathbb{C}^{m\times n}$,

$$\mathrm{rank}(A) + \dim(\mathrm{null}(A)) = n.$$

3 Determinants

If $A = [\alpha] \in \mathbb{C}^{1\times 1}$, then its determinant is given by $\det(A) = \alpha$. The **determinant** of an $n \times n$ matrix A can be defined in terms of order $n-1$ determinants (expansion in cofactors):

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} a_{ij}(-1)^{i+j}\det(A_{ij}) \quad \text{for any } i, \\ &= \sum_{i=1}^{n} a_{ij}(-1)^{i+j}\det(A_{ij}) \quad \text{for any } j. \end{aligned}$$

Here A_{ij} is an $(n-1)\times(n-1)$ submatrix of A obtained by deleting the ith row and jth column. The scalar $(-1)^{i+j}\det A_{ij}$ is called a **cofactor** of A.

Useful properties of the determinant include

$$\det(AB) = \det(A)\det(B), \quad \det(\alpha A) = \alpha^n \det(A) \quad (\alpha \in \mathbb{C}).$$

If A is block diagonal (or block triangular) with square diagonal blocks $A_{11}, A_{22}, \dots, A_{pp}$ (of possibly different sizes) then

$$\det(A) = \det(A_{11})\det(A_{22})\cdots\det(A_{pp}).$$

4 Inverses

If $A, B \in \mathbb{C}^{n\times n}$ satisfy $AB = I$ then B is the **inverse** of A, written $B = A^{-1}$. If A^{-1} exists A is **nonsingular**; otherwise A is **singular**. Also,

$$(AB)^{-1} = B^{-1}A^{-1}, \quad (A^{-1})^T = (A^T)^{-1} = A^{-T}.$$

Theorem 1 *For $A \in \mathbb{C}^{n\times n}$ the following conditions are equivalent to A being nonsingular:*

1. $\mathrm{null}(A) = \{0\}$ *(i.e, there is no nonzero $y \in \mathbb{C}^n$ such that $Ay = 0$).*
2. $\mathrm{rank}(A) = n$ *(i.e., the rows or columns of A are linearly independent).*
3. $\det(A) \neq 0$.
4. *None of A's eigenvalues is zero.*

The **adjugate** adj of matrix is the transpose of the matrix of cofactors:

$$\mathrm{adj}(A)_{ij} = \big((-1)^{i+j} \det A_{ij}\big)^T,$$

where A_{ij} is the submatrix of A obtained by deleting the ith row and jth column. A calculation using the cofactor expansion for the determinant shows that $A\,\mathrm{adj}(A) = \det(A)I$ so that if A is nonsingular ($\det A \neq 0$) then

$$A^{-1} = \frac{\mathrm{adj}(A)}{\det(A)}.$$

Exercises

1. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$, $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, and let $b = Ax$.
(i) Calculate b by using three vector inner products.
(ii) Calculate b by forming a linear combination of columns of A.

2. Let A be the $n \times n$ matrix with $a_{j,j+1} = 1$, $j = 1, \ldots, n-1$ and all other elements zero. Represent A as a sum of vector outer products.

3. (i) Show that

$$A = \begin{bmatrix} \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \end{bmatrix}$$

is orthogonal and hence determine A^{-1}.
(ii) Show that the product of two orthogonal matrices is an orthogonal matrix.
(iii) If Q is a real orthogonal matrix, what is $\det(Q)$? If U is a unitary matrix, what is $\det(U)$?

4. Consider the block matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ with $A \in \mathbb{C}^{n\times n}$ and $D \in \mathbb{C}^{m\times m}$ nonsingular.

(i) Determine $X \in \mathbb{C}^{m\times m}$ and $Y \in \mathbb{C}^{n\times n}$ such that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & O \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ O & X \end{bmatrix} \qquad \text{(block LU factorization)}$$

and that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ O & I \end{bmatrix} \begin{bmatrix} Y & O \\ C & D \end{bmatrix} \qquad \text{(block UL factorization).}$$

(ii) Prove that

$$\det(A)\det(D - CA^{-1}B) = \det(D)\det(A - BD^{-1}C), \tag{1}$$

$$\det(I - CB) = \det(I - BC), \tag{2}$$

$$\det(I - xy^*) = 1 - y^*x, \tag{3}$$

where $x, y \in \mathbb{C}^n$.

5. (i) Verify the **Sherman–Morrison** formula: if $A \in \mathbb{C}^{n\times n}$ is nonsingular and $u, v \in \mathbb{C}^n$ are such that $v^*A^{-1}u \neq -1$, then

$$(A + uv^*)^{-1} = A^{-1} - \frac{A^{-1}uv^*A^{-1}}{1 + v^*A^{-1}u}. \tag{4}$$

(ii) Let $\alpha \in \mathbb{C}$. When is $I + \alpha e_i e_j^T$ nonsingular? Determine the inverse in those cases where it exists.

6. Let $x_1, x_2, \ldots, x_n$ be nonzero orthogonal vectors. Show that $x_1, \ldots, x_n$ are linearly independent. Hint: assume the linearly dependence relation $\sum \alpha_i x_i = 0$ and show that orthogonality implies that all the α_i must be zero.

7. (i) Which is the only matrix that is both idempotent and involutory?

(ii) Which is the only matrix that is both idempotent and nilpotent?

(iii) Let $x, y \in \mathbb{C}^n$. When is xy^* idempotent? When is it nilpotent?

(iv) Prove that if A, B are idempotent and $AB = BA$ then AB is also idempotent.

(v) Prove that an idempotent matrix is singular unless it is the identity matrix.

(vi) Prove that A is involutory if and only if $(I - A)(I + A) = O$.

(vii) Let $x \in \mathbb{C}^n$ and $x^*x = 1$. Show that $I - 2xx^*$ is involutory. The matrix $I - 2xx^*$ is called a **Householder reflector**.

(viii) Prove that if $(I - A)^{-1} = \sum_{j=0}^{k} A^j$ for some integer $k \geq 0$ then A is nilpotent.

8. Let A, B be $n \times n$ upper triangular matrices with diagonal elements a_{jj} and b_{jj}, respectively. Show that

(i) AB is triangular and the diagonal elements of AB are $a_{ii}b_{ii}$.

(ii) If $a_{jj} \neq 0$ for all j then A is nonsingular, A^{-1} is triangular and the diagonal elements of A^{-1} are $1/a_{jj}$.

9. Let $A \in \mathbb{C}^{n\times n}$. The **trace** of A is defined as the sum of its diagonal elements, i.e, $\text{trace}(A) = \sum_{i=1}^{n} a_{ii}$.

(i) Show that the trace is a linear function, i.e., if $A, B \in \mathbb{C}^{n\times n}$ and $\alpha, \beta \in \mathbb{C}$, then

$$\text{trace}(\alpha A + \beta B) = \alpha\,\text{trace}(A) + \beta\,\text{trace}(B).$$

(ii) Show that $\text{trace}(AB) = \text{trace}(BA)$, even though in general $AB \neq BA$.

(iii) Show that if $S \in \mathbb{R}^{n\times n}$ is skew-symmetric, i.e., $S^T = -S$, the $\text{trace}(S) = 0$. Prove the converse to this statement or provide a counterexample.