MA1101R Cheatsheet 17/18 Sem 1 Finals

by Lee Yiyuan and Eugene Lim

Matrices

Definition 2.5.2. Let $\boldsymbol{A} = (a_{ij})$ be an $n \times n$ matrix. Let $\boldsymbol{M}_{ij}$ be an $(n-1) \times (n-1)$ matrix obtained from $\boldsymbol{A}$ by deleting the ith row and the jth column. Then the *determinant* of $\boldsymbol{A}$ is defined as

$$\det(\boldsymbol{A}) = \begin{cases} a_{11} & \text{if } n = 1 \\ a_{11}A_{11} + \cdots + a_{1n}A_{1n} & \text{if } n > 1 \end{cases}$$

where

$$A_{ij} = (-1)^{i+j} \det\left(\boldsymbol{M_{ij}}\right)$$

The number A_{ij} is called the (i, j)*-cofactor* of $\boldsymbol{A}$.

Theorem 2.5.8. The determinant of a triangular matrix is equal to the product of its diagonal entries.

Theorem 2.5.15. Let $\boldsymbol{A}$ be a square matrix.

1. If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by multiplying one row of $\boldsymbol{A}$ by a constant k, then $\det(\boldsymbol{B}) = k\det(\boldsymbol{A})$.
2. If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by interchanging two rows, then $\det(\boldsymbol{B}) = -\det(\boldsymbol{A})$.
3. If $\boldsymbol{B}$ is obtained from $\boldsymbol{A}$ by adding a multiple of one row of $\boldsymbol{A}$ to another row, then $\det(\boldsymbol{B}) = \det(\boldsymbol{A})$.
4. Let $\boldsymbol{E}$ be an elementary matrix of the same size as $\boldsymbol{A}$. Then $\det(\boldsymbol{EA}) = \det(\boldsymbol{E})\det(\boldsymbol{A})$.

Theorem 2.5.25. If $\boldsymbol{A}$ is invertible, then $\boldsymbol{A}^{-1} = \frac{1}{\det(\boldsymbol{A})}\,\mathbf{adj}(\boldsymbol{A})$.

Theorem 2.5.27. Suppose $\boldsymbol{Ax} = \boldsymbol{b}$ is a linear system where $\boldsymbol{A}$ is an $n \times n$ matrix. Let $\boldsymbol{A_i}$ be the matrix obtained from $\boldsymbol{A}$ be replacing the ith column of $\boldsymbol{A}$ by $\boldsymbol{b}$. If $\boldsymbol{A}$ is invertible, then the system has only one solution

$$\boldsymbol{x} = \frac{1}{\det(\boldsymbol{A})} \begin{pmatrix} \det\left(\boldsymbol{A_1}\right) \\ \vdots \\ \det\left(\boldsymbol{A_n}\right) \end{pmatrix}$$

Definition 2.5.24. Let $\boldsymbol{A}$ be a square matrix of order n. Then the *(classical) adjoint* of $\boldsymbol{A}$ is the $n \times n$ matrix

$$\mathbf{adj}(\boldsymbol{A}) = (A_{ij})^T_{n \times n}$$

where A_{ij} is the (i, j)-cofactor of $\boldsymbol{A}$.

Euclidean Spaces

Definition 3.2.3. Let $S = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ be a set of vectors in $\mathbb{R}^n$. Then the set of all linear combinations of $\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}$,

$$\{c_1\boldsymbol{u_1} + \cdots + c_k\boldsymbol{u_k} \mid c_1, \ldots, c_k \in \mathbb{R}\}$$

is called the *linear span* of S (or the *linear span* of $\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}$) and is denoted by $\text{span}(S)$ (or $\text{span}\{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$).

Theorem 3.2.10. Let $S_1 = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ and $S_2 = \{\boldsymbol{v_1}, \ldots, \boldsymbol{v_m}\}$ be subsets of $\mathbb{R}^n$. Then $\text{span}\,(S_1) \subseteq \text{span}\,(S_2)$ if and only if each u_i is a linear combination of $\boldsymbol{v_1}, \ldots, \boldsymbol{v_m}$.

Definition 3.3.2. Let V be a subset of $\mathbb{R}^n$. Then V is called a *subspace* of $\mathbb{R}^n$ if $V = \text{span}(S)$ where $S = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ for some vectors $\boldsymbol{u_1}, \ldots, \boldsymbol{u_k} \in \mathbb{R}^n$.

More precisely, V is called the *subspace spanned* by S (or the *subspace spanned* by $\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}$). We also say that S *spans* (or $\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}$ *span*) the subspace V.

Remark 3.3.8. Let V be a non-empty subset of $\mathbb{R}^n$. Then V is a subspace of $\mathbb{R}^n$ if and only if

$$\text{for all } \boldsymbol{u}, \boldsymbol{v} \in V \text{ and } c, d \in \mathbb{R},\ \ c\boldsymbol{u} + d\boldsymbol{v} \in V$$

Definition 3.4.2. Let $S = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ be a set of vectors in $\mathbb{R}^n$. Consider the equation

$$c_1\boldsymbol{u_1} + \cdots + c_k\boldsymbol{u_k} = \boldsymbol{0}$$

where $c_1, \ldots, c_k$ are variables.

1. S is called a *linearly dependent set* and $\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}$ are said to be *linearly independent* if the equation has only the trivial solution $c_1 = \cdots = c_k = 0$.
2. S is called a *linearly independent set* and $\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}$ are said to be *linearly dependent* if the equation has non-trivial solutions.

Definition 3.5.4. Let $S = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ be a subset of a vector space V. Then S is called a *basis* for V if S is linearly independent and S spans V.

Definition 3.5.8. Let $S = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ be a basis for a vector space V and $\boldsymbol{v}$ a vector in V. By Theorem 3.5.7, $\boldsymbol{v}$ is expressed uniquely as a linear combination

$$v = c_1\boldsymbol{u_1} + \cdots + c_k\boldsymbol{u_k}$$

The coefficients $c_1, \ldots, c_k$ are called the *coordinates* of $\boldsymbol{v}$ relative to the basis S.

The vector $(\boldsymbol{v})_S = (c_1, \ldots, c_k) \in \mathbb{R}^k$ is called the *coordinate vector* of $\boldsymbol{v}$ relative to the basis S.

Theorem 3.6.1. Let V be a vector space which has a basis with k vectors. Then

1. any subset of V with more than k vectors is always linearly dependent;
2. any subset of V with less than k vectors cannot span V.

Definition 3.6.3. The *dimension* of a vector space V, denoted by $\dim(V)$, is defined to be the number of vectors in a basis for V. In addition, we define the dimension of the zero space to be zero.

Theorem 3.6.7. Let V be a vector space of dimension k and S a subset of V. The following are equivalent:

1. S is a basis for V.
2. S is linearly independent and $|S| = k$.
3. S spans V and $|S| = k$.

Definition 3.7.3. Let $S = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ and T be two bases for a vector space. The square matrix $\boldsymbol{P} = \begin{pmatrix}[\boldsymbol{u_1}]_T & \cdots & [\boldsymbol{u_2}]_T\end{pmatrix}$ is called the *transition matrix* from S to T.

Vector Space of Matrices

Definition 4.1.2. Let $\boldsymbol{A} = (a_{ij})$ be an $m \times n$ matrix. The *row space* of $\boldsymbol{A}$ is the subspace of $\mathbb{R}^n$ spanned by the rows of $\boldsymbol{A}$. The *column space* of $\boldsymbol{A}$ is the subspace of $\mathbb{R}^m$ spanned by the columns of $\boldsymbol{A}$.

Theorem 4.1.7. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be row equivalent matrices. Then the row space of $\boldsymbol{A}$ and the row space of $\boldsymbol{B}$ are identical, i.e. elementary row operations preserve the row space of a matrix.

Theorem 4.1.11. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be row equivalent matrices. Then the following statements hold:

1. A given set of columns of $\boldsymbol{A}$ is linearly independent if and only if the set of corresponding columns of $\boldsymbol{B}$ is linearly independent.
2. A given set of columns of $\boldsymbol{A}$ forms a basis for the column space of $\boldsymbol{A}$ if and only if the set of corresponding columns of $\boldsymbol{B}$ forms a basis for the column space of $\boldsymbol{B}$.

Theorem 4.2.1. The row space and column space of a matrix have the same dimension.

Definition 4.2.3. The *rank* of a matrix is the dimension of its row space (or column space). We denote the rank of a matrix $\boldsymbol{A}$ by $\text{rank}(\boldsymbol{A})$. Note that $\text{rank}(\boldsymbol{A})$ is equal to the number of nonzero rows as well as the number of pivot columns in a row-echelon form of $\boldsymbol{A}$.

Theorem 4.2.8. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $m \times n$ and $n \times p$ matrices respectively. Then

$$\text{rank}(\boldsymbol{AB}) \leq \min\{\text{rank}(\boldsymbol{A}), \text{rank}(\boldsymbol{B})\}$$

Definition 4.3.1. Let $\boldsymbol{A}$ be an $m \times n$ matrix. The solution space of the homogeneous system of linear equations $\boldsymbol{Ax} = \boldsymbol{0}$ is known as the *nullspace* of $\boldsymbol{A}$.

The dimension of the null space of a matrix $\boldsymbol{A}$ is known as the *nullity* of $\boldsymbol{A}$ and is denoted by $\text{nullity}(\boldsymbol{A})$. If $\boldsymbol{A}$ is an $m \times n$ matrix, it is clear that $\text{nullity}(\boldsymbol{A}) \leq n$ since the nullspace is a subspace of $\mathbb{R}^n$.

Theorem 4.3.4. Let $\boldsymbol{A}$ be a matrix with n columns. Then

$$\text{rank}(\boldsymbol{A}) + \text{nullity}(\boldsymbol{A}) = n$$

Theorem 4.3.6. Suppose the linear equations $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution $\boldsymbol{v}$. Then the solution set of the system is given by

$$M = \{\boldsymbol{u} + \boldsymbol{v} \mid \boldsymbol{u} \text{ is an element of the nullspace of } \boldsymbol{A}\}$$

Orthogonality

Definition 5.2.1.

1. Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^n$ are called *orthogonal* if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.
2. A set S of vectors in $\mathbb{R}^n$ is called *orthogonal* if every pair of distinct vectors in S are orthogonal.
3. A set S of vectors in $\mathbb{R}^n$ is called *orthonormal* if S is orthogonal and every vector in S is a unit vector.

Definition 5.2.4.

1. A basis S for a vector space is called an *orthogonal basis* if S is orthogonal.
2. A basis S for a vector space is called an *orthonormal basis* if S is orthonormal.

Theorem 5.2.8. If $S = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ is an orthogonal basis for a vector space V, then for any vector $\boldsymbol{w}$ in V,

$$\boldsymbol{w} = \frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}\boldsymbol{u_1} + \cdots + \frac{\boldsymbol{w} \cdot \boldsymbol{u_k}}{\boldsymbol{u_k} \cdot \boldsymbol{u_k}}\boldsymbol{u_k}$$

i.e. $(\boldsymbol{w})_S = \left(\frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}, \ldots, \frac{\boldsymbol{w} \cdot \boldsymbol{u_k}}{\boldsymbol{u_k} \cdot \boldsymbol{u_k}}\right)$

Definition 5.2.10. Let V be a subspace of $\mathbb{R}^n$. A vector $\boldsymbol{u} \in \mathbb{R}^n$ is said to be *orthogonal* (or *perpendicular*) to V if $\boldsymbol{u}$ is orthogonal to all vectors in V.

Definition 5.2.13. Let V be a subspace of $\mathbb{R}^n$. Every vector $\boldsymbol{u} \in \mathbb{R}^n$ can be written uniquely as

$$\boldsymbol{u} = \boldsymbol{n} + \boldsymbol{p}$$

such that $\boldsymbol{n}$ is a vector orthogonal to V and $\boldsymbol{p}$ is a vector in V. The vector $\boldsymbol{p}$ is called the (*orthogonal*) *projection* of $\boldsymbol{u}$ onto V.

Theorem 5.2.15. Let V be a subspace of $\mathbb{R}^n$ and $\boldsymbol{w}$ a vector in $\mathbb{R}^n$. If $\{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ is an orthogonal basis for V, then

$$\frac{\boldsymbol{w} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}\boldsymbol{u_1} + \cdots + \frac{\boldsymbol{w} \cdot \boldsymbol{u_k}}{\boldsymbol{u_k} \cdot \boldsymbol{u_k}}\boldsymbol{u_k}$$

is the projection of $\boldsymbol{w}$ onto V.

Theorem 5.2.19. Let $\{\boldsymbol{u_1}, \ldots, \boldsymbol{u_k}\}$ be a basis for a vector space V. Let

$$\begin{aligned}
\boldsymbol{v_1} &= \boldsymbol{u_1} \\
\boldsymbol{v_2} &= \boldsymbol{u_2} - \frac{\boldsymbol{u_2} \cdot \boldsymbol{v_1}}{\boldsymbol{v_1} \cdot \boldsymbol{v_1}}\boldsymbol{v_1} \\
&\vdots \\
\boldsymbol{v_k} &= \boldsymbol{u_k} - \frac{\boldsymbol{u_k} \cdot \boldsymbol{v_1}}{\boldsymbol{v_1} \cdot \boldsymbol{v_1}}\boldsymbol{v_1} - \cdots - \frac{\boldsymbol{u_k} \cdot \boldsymbol{v_{k-1}}}{\boldsymbol{v_{k-1}} \cdot \boldsymbol{v_{k-1}}}\boldsymbol{v_{k-1}}
\end{aligned}$$

Note that the right side is the projection of the vectors onto the orthogonal basis. To convert it into an orthonormal basis for V, simply divide each $\boldsymbol{v_i}$ by their length.

Theorem 5.3.10. Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a linear system. Then $\boldsymbol{u}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$ if and only if $\boldsymbol{u}$ is a solution to $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$.

Definition 5.4.3. A square matrix $\boldsymbol{A}$ is called *orthogonal* if $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$.

Eigens and Diagonalization

Definition 6.1.3. Let $\boldsymbol{A}$ be a square matrix of order n. A nonzero column vector $\boldsymbol{u}$ in $\mathbb{R}^n$ is called an *eigenvector* of $\boldsymbol{A}$ if

$$\boldsymbol{Au} = \lambda\boldsymbol{u}$$

for some scalar λ. The scalar λ is called an *eigenvalue* of $\boldsymbol{A}$ and $\boldsymbol{u}$ is said to be an eigenvector of $\boldsymbol{A}$ *associated* with the eigenvalue λ.

Definition 6.1.6. Let $\boldsymbol{A}$ be a square matrix of order n. The equation

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$$

is called the *characteristic equation* of $\boldsymbol{A}$ and the polynomial

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A})$$

is called the *characteristic polynomial* of $\boldsymbol{A}$.

Theorem 6.1.8. Let $\boldsymbol{A}$ be an $n \times n$ matrix. The following statements are equivalent:

1. $\boldsymbol{A}$ is invertible.
2. The linear system $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution.
3. The reduced row-echelon form of $\boldsymbol{A}$ is an identity matrix.
4. $\boldsymbol{A}$ can be expressed as a product of elementary matrices.
5. $\det(\boldsymbol{A}) \neq 0$
6. The rows of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$.
7. The columns of $\boldsymbol{A}$ form a basis for $\mathbb{R}^n$.
8. $\operatorname{rank}(\boldsymbol{A}) = n$
9. 0 is not an eigenvalue of $\boldsymbol{A}$.

Theorem 6.1.9. If $\boldsymbol{A}$ is a triangular matrix, the eigenvalues of $\boldsymbol{A}$ are the diagonal entries of $\boldsymbol{A}$.

Definition 6.1.11. Let $\boldsymbol{A}$ be a square matrix of order n and λ an eigenvalue of $\boldsymbol{A}$. Then the solution space of the linear system $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is called the *eigenspace* of $\boldsymbol{A}$ *associated* with the eigenvalue λ and is denoted by E_λ.

Note that if $\boldsymbol{u}$ is a nonzero vector in E_λ, then $\boldsymbol{u}$ is an eigenvector of $\boldsymbol{A}$ associated with the eigenvalue λ.

Definition 6.2.1. A square matrix $\boldsymbol{A}$ is called *diagonalizable* if there exists an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{AP}$ is a diagonal matrix. Here the matrix $\boldsymbol{P}$ is said to *diagonalize* $\boldsymbol{A}$.

Theorem 6.2.3. Let $\boldsymbol{A}$ be a square matrix of order n. Then $\boldsymbol{A}$ is diagonlizable if and only if $\boldsymbol{A}$ has n linearly independent eigenvectors.

Remark 6.2.5.2. The dimension of an eigenspace E_λ of a square matrix $\boldsymbol{A}$ associated with the eigenvalue λ is at most the multiplicity of λ in the characteristic polynomial of $\boldsymbol{A}$.

Furthermore, $\boldsymbol{A}$ is diagonalizable if and only if the dimension of each eigenspace of $\boldsymbol{A}$ is equal to the multiplicity of its associated eigenvalue.

Theorem 6.2.7. Let $\boldsymbol{A}$ be a square matrix of order n. If $\boldsymbol{A}$ has n distinct eigenvalues, then $\boldsymbol{A}$ is diagonalizable.

Definition 6.3.2. A square matrix $\boldsymbol{A}$ is called *orthogonally diagonalizable* if there exists an orthogonal matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^T\boldsymbol{AP}$ is a diagonal matrix. Here the matrix $\boldsymbol{P}$ is said to *orthogonally diagonalize* $\boldsymbol{A}$.

Theorem 6.3.4. A square matrix is orthogonally diagonalizable if and only if it is symmetric.

Linear Transformations

Definition 7.1.1. A *linear transformation* is a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ of the form

$$T\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

where $a_{11}, \ldots, a_{mn}$ are real numbers. In particular, if $n = m$, T is also called a *linear operator* on $\mathbb{R}^n$. We can rewrite the formula of T as

$$T\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11} \cdots a_{1n} \\ \vdots \\ a_{m1} \cdots a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

The matrix $(a_{ij})_{m \times n}$ above is called the *standard matrix* for T.

Definition 7.1.10. Let $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^m \to \mathbb{R}^k$ be linear transformations. The *composition* of T with S, denoted by $T \circ S$, is a mapping from $\mathbb{R}^n \to \mathbb{R}^k$ such that

$$(T \circ S)(\boldsymbol{u}) = T(S(\boldsymbol{u})) \text{ for } \boldsymbol{u} \in \mathbb{R}^n$$

Definition 7.2.1. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The *range* of T, denoted by $R(T)$, is the set of images of T, i.e.

$$R(T) = \{T(\boldsymbol{u}) \mid \boldsymbol{u} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

Theorem 7.2.4. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $\boldsymbol{A}$ the standard matrix for T. Then the range of T is defined as:

$$R(T) = \text{the column space of } \boldsymbol{A}$$

which is a subspace of $\mathbb{R}^m$. This is also called the range of the linear transformation.

Definition 7.2.5. Let T be a linear transformation. The dimension of $R(T)$ is called the *rank* of T and is denoted by $\operatorname{rank}(T)$.

By Theorem 7.2.4, if $\boldsymbol{A}$ is the standard matrix for T, then $\operatorname{rank}(T) = \operatorname{rank}(\boldsymbol{A})$.

Definition 7.2.7. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The *kernel* of T, denoted by $\operatorname{Ker}(T)$, is the set of vectors in $\mathbb{R}^n$ whose image is the zero vector in $\mathbb{R}^m$, i.e.

$$\operatorname{Ker}(T) = \{\boldsymbol{u} \mid T(\boldsymbol{u}) = \boldsymbol{0}\} \subseteq \mathbb{R}^n$$

This is also called the nullity of T.

Theorem 7.2.9. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $\boldsymbol{A}$ the standard matrix for T. Then

$$\operatorname{Ker}(T) = \text{the nullspace of } \boldsymbol{A}$$

Definition 7.2.10. Let T be a linear transformation. The dimension of $\operatorname{Ker}(T)$ is called the *nullity* of T and is denoted by $\operatorname{nullity}(T)$.

By Theorem 7.2.9, if $\boldsymbol{A}$ is the standard matrix for T, then $\operatorname{nullity}(T) = \operatorname{nullity}(\boldsymbol{A})$.

Theorem 7.2.13. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then

$$\operatorname{rank}(T) + \operatorname{nullity}(T) = n$$