2 Review of Linear Algebra and Matrices

2.1 Vector Spaces

2.1 Definition: A (real) vector space consists of a non empty set $\mathbf{V}$ of elements called vectors, and two operations:

(1) Addition is defined for pairs of elements in $\mathbf{V}$, $\mathbf{x}$ and $\mathbf{y}$, and yields an element in $\mathbf{V}$, denoted by $\mathbf{x}+\mathbf{y}$.

(2) Scalar multiplication, is defined for the pair α, a real number, and an element $\mathbf{x} \in \mathbf{V}$, and yields an element in $\mathbf{V}$ denoted by $\alpha\mathbf{x}$.

Eight properties are assumed to hold for $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{V}, \alpha, \beta, 1 \in \mathbb{R}$:

(1) $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$

(2) $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$

(3) There is an element in $\mathbf{V}$ denoted $\mathbf{0}$ such that $\mathbf{0}+\mathbf{x} = \mathbf{x}+\mathbf{0} = \mathbf{x}$

(4) For each $\mathbf{x} \in \mathbf{V}$ there is an element in $\mathbf{V}$ denoted $-\mathbf{x}$ such that $\mathbf{x}+(-\mathbf{x}) = (-\mathbf{x})+\mathbf{x} = \mathbf{0}$

(5) $\alpha(\mathbf{x}+\mathbf{y}) = \alpha\mathbf{x}+\alpha\mathbf{y}$ for all α

(6) $(\alpha+\beta)\mathbf{x} = \alpha\mathbf{x}+\beta\mathbf{x}$ for all α, β

(7) $1\mathbf{x} = \mathbf{x}$

(8) $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$ for all α, β

2.2 Definition: Vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are linearly independent if $\sum_i c_i\mathbf{a}_i \neq 0$ unless $c_i = 0$ for all i.

2.3 Definition: A linear basis or coordinate system in a vector space $\mathbf{V}$ is a set $\mathbf{B}$ of linearly independent vectors in $\mathbf{V}$ such that each vector in $\mathbf{V}$ can be written as a linear combination of the vectors in $\mathbf{B}$.

2.4 Definition: The dimension of a vector space is the number of vectors in any basis of the vector space.

2.5 Definition: Let $\mathbf{V}$ be a p dimensional vector space and let $\mathbf{W}$ be an n dimensional vector space. A linear transformation $\mathbf{L}$ from $\mathbf{V}$ to $\mathbf{W}$ is a mapping (function) from $\mathbf{V}$ to $\mathbf{W}$ such that

$$\mathbf{L}(\alpha\mathbf{x}+\beta\mathbf{y}) = \alpha\mathbf{L}(\mathbf{x})+\beta\mathbf{L}(\mathbf{y}) \quad \text{for every} \quad \mathbf{x}, \mathbf{y} \in \mathbf{V} \quad \text{and all} \quad \alpha, \beta \in \mathbb{R}.$$

2.2 Vectors and Matrices

2.2.1 Matrix Notation and Elementary Properties

2.6 Definition: Matrix: An $m \times n$ matrix with elements a_{ij} is denoted $\mathbf{A} = (a_{ij})_{m\times n}$.

2.7 Definition: Vector: A vector of length n is denoted $\mathbf{a} = (a_i)_n$. If all elements equal 1 it is denoted $\mathbf{1}_n$.

2.8 Definition: Diagonal Matrix:

$$\mathrm{diag}(a_1, \ldots, a_n) \equiv \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{pmatrix}.$$

2.9 Definition: Identity Matrix: $\mathbf{I}_{n\times n} \equiv \mathrm{diag}(\mathbf{1}_n)$.

2.10 Definition: Matrix Transpose: If $\mathbf{A} = (a_{ij})_{m\times n}$, then $\mathbf{A}' \equiv (a'_{ij})_{n\times m}$ where $a'_{ij} = a_{ji}$.

2.11 Definition: If $\mathbf{A} = \mathbf{A}'$, then $\mathbf{A}$ is symmetric.

2.12 Definition: Matrix Sum: If $\mathbf{A} = (a_{ij})_{m\times n}$ and $\mathbf{B} = (b_{ij})_{m\times n}$, then $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})_{m\times n}$.

2.13 Theorem: Matrix sums satisfy $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$.

2.14 Definition: Matrix Product: If $\mathbf{A} = (a_{ij})_{m\times n}$ and $\mathbf{B} = (b_{ij})_{n\times p}$, then $\mathbf{AB} = (c_{ij})_{m\times p}$, where $c_{ij} = \sum_k a_{ik} b_{kj}$.

2.15 Theorem: Matrix products satisfy $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.

2.16 Definition: Matrix Trace: The sum of the diagonal elements, $\mathrm{tr}(\mathbf{A}) \equiv \sum_i a_{ii}$.

2.17 Theorem: The trace satisfies $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$ if $\mathbf{A} = (a_{ij})_{m\times n}$ and $\mathbf{B} = (b_{ij})_{m\times n}$, and $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ if $\mathbf{A}$ and $\mathbf{B}$ are square matrices.

2.2.2 Range, Rank, and Null Space

2.18 Definition: Range (Column Space): $\mathcal{R}(\mathbf{A}) \equiv$ the linear space spanned by the columns of $\mathbf{A}$.

2.19 Definition: Rank: $\mathrm{rank}(\mathbf{A}) \equiv \mathrm{r}(\mathbf{A}) \equiv$ the number of linearly independent columns of $\mathbf{A}$ (i.e., the dimension of $\mathcal{R}(\mathbf{A})$), or equivalently, the number of linearly independent rows of $\mathbf{A}$.

2.20 Theorem: Decreasing property of rank: $\mathrm{rank}(\mathbf{AB}) \le \min\{\mathrm{rank}(\mathbf{A}), \mathrm{rank}(\mathbf{B})\}$.

2.21 Definition: Null Space: $\mathcal{N}(\mathbf{A}) \equiv \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{0}\}$. The nullity of $\mathbf{A}$ is the dimension of $\mathcal{N}(\mathbf{A})$.

2.22 Theorem: $\mathrm{rank}(\mathbf{A}) + \mathrm{nullity}(\mathbf{A}) = n$, the number of columns of $\mathbf{A}$.

2.23 Theorem: $\mathrm{r}(\mathbf{A}) = \mathrm{r}(\mathbf{A}') = \mathrm{r}(\mathbf{A}'\mathbf{A}) = \mathrm{r}(\mathbf{A}\mathbf{A}')$.

2.2.3 Inverse

2.24 Definition: An $n \times n$ matrix $\mathbf{A}$ is invertible (or non-singular) if there is a matrix $\mathbf{A}^{-1}$ such that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_{n\times n}$. Equivalently, $\mathbf{A}$ $(n \times n)$ is invertible if and only if $\mathrm{rank}(\mathbf{A}) = n$.

2.25 Theorem: Inverse of Product: $(\mathbf{AB})^{-\mathbf{1}} = \mathbf{B}^{-\mathbf{1}}\mathbf{A}^{-\mathbf{1}}$ if $\mathbf{A}$ and $\mathbf{B}$ are invertible.

2.2.4 Inner Product, Length, and Orthogonality

2.26 Definition: Inner product: $\mathbf{a}'\mathbf{b} = \sum_i a_i b_i$, where $\mathbf{a} = (a_i)$, $\mathbf{b} = (b_i)$.

2.27 Definition: Vector norm (length): $||\mathbf{a}|| = \sqrt{\mathbf{a}'\mathbf{a}}$.

2.28 Definition: Orthogonal vectors: $\mathbf{a} = (a_i)$ and $\mathbf{b} = (b_i)$ are orthogonal if $\mathbf{a}'\mathbf{b} = 0$.

2.29 Definition: Orthogonal matrix: $\mathbf{A}$ is orthogonal if its columns are orthogonal vectors of length 1, or equivalently, if $\mathbf{A}^{-1} = \mathbf{A}'$.

2.2.5 Determinants

2.30 Definition: For a square matrix $\mathbf{A}$, $|\mathbf{A}| \equiv \sum_i a_{ij} A_{ij}$, where the cofactor $A_{ij} = (-1)^{i+j}|\mathbf{M}_{ij}|$, and $\mathbf{M}_{ij}$ is the matrix obtained by deleting the ith row and jth column from $\mathbf{A}$.

2.31 Theorem: $\left|\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right| = ad - bc.$

2.32 Theorem: $|\mathbf{A}| = 0,$ if and only if $\mathbf{A}$ is singular.

2.33 Theorem: $|\mathrm{diag}(a_1, \ldots, a_n)| = \prod_i a_i$.

2.34 Theorem: $|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|$.

2.35 Theorem: $\left|\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}\right| = |\mathbf{A}| \cdot |\mathbf{C}|.$

2.2.6 Eigenvalues

2.36 Definition: If $\mathbf{Ax} = \lambda\mathbf{x}$ where $\mathbf{x} \neq 0$, then λ is an eigenvalue of $\mathbf{A}$ and $\mathbf{x}$ is a corresponding eigenvector.

Let $\mathbf{A}$ be a symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$.

2.37 Theorem: (Spectral Theorem, a.k.a. Principal Axis Theorem) For any symmetric matrix $\mathbf{A}$ there exists an orthogonal matrix $\mathbf{T}$ such that: $\mathbf{T}'\mathbf{AT} = \mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

2.38 Theorem: $\mathrm{r}(\mathbf{A}) =$ the number of non-zero λ_i

2.39 Theorem: $\mathrm{tr}(\mathbf{A}) = \sum_i \lambda_i$.

2.40 Theorem: $|\mathbf{A}| = \prod_i \lambda_i$.

2.2.7 Positive Definite and Semidefinite Matrices

2.41 Definition: A symmetric matrix $\mathbf{A}$ is positive semidefinite (p.s.d.) if $\mathbf{x}'\mathbf{Ax} \geq 0$ for all $\mathbf{x}$.

Properties of a p.s.d matrix $\mathbf{A}$:

2.42 Theorem: The diagonal elements a_{ii} are all non-negative.

2.43 Theorem: All eigenvalues of $\mathbf{A}$ are nonnegative.

2.44 Theorem: $\mathrm{tr}(\mathbf{A}) \geq 0$.

2.45 Definition: A symmetric matrix $\mathbf{A}$ is called positive definite (p.d.) if $\mathbf{x}'\mathbf{Ax} > 0$ for all non-zero $\mathbf{x}$.

Properties of a p.d matrix $\mathbf{A}$:

2.46 Theorem: All diagonal elements and all eigenvalues of $\mathbf{A}$ are positive.

2.47 Theorem: $\mathrm{tr}(\mathbf{A}) > 0$.

2.48 Theorem: $|\mathbf{A}| > 0$.

2.49 Theorem: There is a nonsingular $\mathbf{R}$ such that $\mathbf{A} = \mathbf{RR}'$ (necessary and sufficient for $\mathbf{A}$ to be p.d.).

2.50 Theorem: $\mathbf{A}^{-1}$ is p.d.

2.2.8 Idempotent and Projection Matrices

2.51 Definition: A matrix $\mathbf{P}$ is idempotent if $\mathbf{P}^2 = \mathbf{P}$. A symmetric idempotent matrix is called a projection matrix.

Properties of a projection matrix $\mathbf{P}$:

2.52 Theorem: If $\mathbf{P}$ is an $n \times n$ matrix and $\text{rank}(\mathbf{P}) = r$, then $\mathbf{P}$ has r eigenvalues equal to 1 and $n - r$ eigenvalues equal to 0.

2.53 Theorem: $\text{tr}(\mathbf{P}) = \text{rank}(\mathbf{P})$.

2.54 Theorem: $\mathbf{P}$ is positive semidefinite.

2.3 Projections

2.55 Definition: For two vectors $\mathbf{x}$ and $\mathbf{y}$, the projection of $\mathbf{y}$ onto $\mathbf{x}$ is

$$\text{Proj}_{\mathbf{x}}(\mathbf{y}) = \frac{\mathbf{x}'\mathbf{y}}{\mathbf{x}'\mathbf{x}}\mathbf{x}.$$

2.56 Theorem: If V is a vector space and Ω is a subspace of V, then $\exists$ two vectors, $\mathbf{w}_1, \mathbf{w}_2 \in V$ such that

1. $\mathbf{y} = \mathbf{w}_1 + \mathbf{w}_2 \quad \forall\, \mathbf{y} \in V$,
2. $\mathbf{w}_1 \in \Omega$ and $\mathbf{w}_2 \in \Omega^{\perp}$.

2.57 Theorem: $\|\mathbf{y} - \mathbf{w}_1\| \leq \|\mathbf{y} - \mathbf{x}\|$ for any $\mathbf{x} \in \Omega$. $\mathbf{w}_1$ is called the projection of $\mathbf{y}$ onto Ω.

2.58 Definition: The matrix $\mathbf{P}$ that takes $\mathbf{y}$ onto $\mathbf{w}_1$ (i.e., $\mathbf{P}\mathbf{y} = \mathbf{w}_1$) is called a projection matrix.

2.59 Theorem: $\mathbf{P}$ projects $\mathbf{y}$ onto the space spanned by the column vectors of $\mathbf{P}$.

2.60 Theorem: $\mathbf{P}$ is a linear transformation.

2.61 Theorem: $\mathbf{I} - \mathbf{P}$ is a projection operator onto $\Omega^{\perp}$.