Review Sheet for Midterm Exam

Terminology

Can you explain these words?

1. reduced row-echelon form
2. elementary row operations
3. consistent
4. rank
5. linear combination
6. linear transformation
7. orthogonal projection
8. image
9. span
10. kernel
11. linear space
12. subspace
13. redundant
14. linearly independent
15. basis
16. trivial relation
17. dimension
18. nullity

19. coordinate vector
20. the matrix of a linear transformation
21. similar matrices
22. isomorophism and isomorphic
23. change of basis matrix
24. orthogonal
25. norm
26. orthonormal
27. Kronecker delta
28. orthogonal complement
29. Gram-Schmidt process
30. QR factorization
31. orthogonal transformation and orthogonal matrix
32. transpose
33. symmetric and skew-symmetric
34. least-squares solution
35. normal equation
36. Moore-Penrose pseudoinverse
37. inner product and inner product space
38. Fourier coefficients

Reduced row-echelon form: If a row has nonzero entries, then the first nonzero entry is 1 (the leading 1). If a column contains a leading 1, then all the other entries in that column are 0. If a row contains a leading 1, then each row above it contains a leading 1 further to the left. The reduced row-echelon form of a matrix A is expressed as $\mathrm{rref}(A)$.

Elementary row operations: To divide a row by a nonzero scalar, to subtract a multiple of a row from another row, and to swap two rows.

Consistent: a system of equations is said to be consistent if there is at least one solution. A system is inconsistent if there is no solution.

Rank: The number of leading 1's in the reduced row-echelon form.

Linear combination: A vector $\vec{b}$ is a linear combination of $\vec{v}_1, \dots, \vec{v}_m$ if $\vec{b}$ is given by $\vec{b} = x_1\vec{v}_1 + \cdots + x_m\vec{v}_m$ with scalars $x_1, \dots, x_m$.

Linear transformation: A transformation T is a linear transformation if $T(\alpha\vec{v} + \beta\vec{w}) = \alpha T(\vec{v}) + \beta T(\vec{w})$ is satisfied for scalars α, β and vectors $\vec{v}$, $\vec{w}$. Then, there exists a matrix A such that $T(\vec{x}) = A\vec{x}$.

Orthogonal projection: For given vector $\vec{x}$ and subspace V, we uniquely have $\vec{x} = \vec{x}_\parallel + \vec{x}_\perp$, where $\vec{x}_\parallel \in V$. We call $\vec{x}_\parallel$ denoted by $\mathrm{proj}_V(\vec{x})$ the orthogonal projection of $\vec{x}$ onto V.

Image: For an $n \times m$ matrix A, $\mathrm{imaga}(A) = \{\vec{y} \in \mathbb{R}^n : \vec{y} = A\vec{x} \text{ for } \forall\vec{x} \in \mathbb{R}^m\}$.

Span: For vectors $\vec{v}_1, \dots, \vec{v}_m \in \mathbb{R}^n$, $\mathrm{span}(\vec{v}_1, \dots, \vec{v}_m) = \{\vec{w} : c_1\vec{v}_1 + \cdots + c_m\vec{v}_m = \vec{w} \text{ with } c_1, \dots, c_m \in \mathbb{R}\}$.

Kernel: For an $n \times m$ matrix A, $\ker(A) = \{\vec{x} \in \mathbb{R}^m : A\vec{x} = \vec{0}\}$.

Linear space: A linear space (vector space) V is a set endowed with a rule for addition and a rule for scalar multiplication such that these operations satisfy the following eight rules ($\vec{x}$, $\vec{y}$, $\vec{z} \in V$ and α, $\beta \in \mathbb{R}$): (i) $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$. (ii) $\vec{x} + \vec{y} = \vec{y} + \vec{x}$. (iii) There exists $\vec{0}$ such that $\vec{x} + \vec{0} = \vec{x}$ for $\forall\vec{x} \in V$. (iv) For each $\vec{x} \in V$, there exists $-\vec{x}$ such that $\vec{x} + (-\vec{x}) = \vec{0}$. (v)

$\alpha(\vec{x}+\vec{y}) = \alpha\vec{x}+\beta\vec{y}$. (vi) $(\alpha+\beta)\vec{x} = \alpha\vec{x}+\beta\vec{x}$. (vii) $\alpha(\beta\vec{x}) = (\alpha\beta)\vec{x}$. (viii) $1\vec{x} = \vec{x}$.

Subspace: A subset W of V is called a subspace of V if (i) W contains $\vec{0} \in V$, (ii) W is closed under addition, and (iii) W is closed under scalar multiplication.

Trivial relation: For vectors $\vec{v}_1, \ldots, \vec{v}_m \in \mathbb{R}^n$, we call $c_1\vec{v}_1+\cdots+c_m\vec{v}_m = \vec{0}$ a (linear) relation. When we choose $c_1 = \cdots = c_m = 0$, the relation is called the trivial relation.

Dimension: The number of vectors in a basis of a subspace V of $\mathbb{R}^n$ is called the dimension of V, denoted by $\dim(V)$.

Nullity: The nullity of matrix A is the dimension of $\ker(A)$.

Coordinate vector: Consider a basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_m)$ of a subspace V of $\mathbb{R}^n$. Any vector $\vec{x} \in V$ can be written uniquely as $\vec{x} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$. The scalars $c_1, \ldots, c_m$ are called the $\mathcal{B}$-coordinates of $\vec{x}$, and the vector $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}$ is called the $\mathcal{B}$-coordinate vector of $\vec{x}$.

The matrix of a linear transformation: Consider a linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ and a basis $\mathcal{B}$ of $\mathbb{R}^n$. The $n \times n$ matrix B such that $[T(\vec{x})]_{\mathcal{B}} = B\,[\vec{x}]_{\mathcal{B}}$ for $\forall \vec{x} \in \mathbb{R}^n$ is called the $\mathcal{B}$-matrix of T.

Similar matrices: An $n \times n$ matrix A is similar to an $n \times n$ matrix B if there exists an invertible matrix S such that $AS = SB$ or $B = S^{-1}AS$.

Isomorophism and isomorphic: An invertible linear transformation T is called an isomorphism. We say the linear space V is isomorphic to the linear space W if there exists an isomorphism T from V to W.

Change of basis matrix: For two bases $\mathcal{A}$ and $\mathcal{B}$ of an n-dimensional linear space V, consider the linear transformation $L_{\mathcal{A}} \circ L_{\mathcal{B}}^{-1}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$, or $S_{\mathcal{B}\to\mathcal{A}}\vec{x} = L_{\mathcal{A}}\left(L_{\mathcal{B}}^{-1}(\vec{x})\right)$ for $\forall \vec{x} \in \mathbb{R}^n$. This invertible matrix $S_{\mathcal{B}\to\mathcal{A}}$ is called the change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$.

Orthogonal: $\vec{v} \cdot \vec{w} = 0$ or more generally $\langle f, g \rangle = 0$. If $\vec{v} \cdot \vec{w} = 0$ for all $\vec{v} \in V$ and $\vec{w} \in W$ for two subspaces V and W of the same linear space, then V and W are said to be orthogonal.

Norm: $||\vec{v}|| = \sqrt{\vec{v} \cdot \vec{v}}$ or more generally $||\vec{v}|| = \sqrt{\langle f, g \rangle}$.

Orthonormal: The vectors are called orthonormal if they are all unit vectors and orthogonal to one another.

Kronecker delta: $\delta_{ij} = 0$ if $i \neq j$ and $= 1$ if $i = j$.

Orthogonal complement: The orthogonal complement $V^\perp$ of a subspace V of $\mathbb{R}^n$ is the set of those vectors $\vec{x} \in \mathbb{R}^n$ that are orthogonal to all vectors in V:

$$V^\perp = \{\vec{x} \in \mathbb{R}^n : \vec{v} \cdot \vec{x} = 0 \text{ for } \forall \vec{v} \in V\}.$$

Note that $V^\perp$ is the kernel of the orthogonal projection onto V.

Gram-Schmidt process: The Gram-Schmidt process is a method of constructing orthonormal vectors from a set of linearly independent vectors (see Theorem 5.2.1).

QR **factorization:** Suppose columns of an $n \times m$ matrix M are linearly independent. Then we can uniquely decompose M as $M = QR$, where the columns of the $n \times m$ matrix Q are orthonormal and R is an $m \times m$ upper triangular matrix with positive diagonal entries.

Orthogonal transformation and orthogonal matrix: A linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ is called orthogonal if it preserves the lengths of vectors: $||T(\vec{x})|| = ||\vec{x}||$ for $\forall \vec{x} \in \mathbb{R}^n$. If $T(\vec{x}) = A\vec{x}$ is an orthogonal transformation, the matrix A is called an orthogonal matrix.

Transpose: Consider an $m \times n$ matrix A. The transpose A^T of A is the $n \times m$ matrix whose ijth entry is the jith entry of A.

Symmetric and skew-symmetric: We say that a square matrix A is symmetric if $A^T = A$, and skew-symmetric if $A^T = -A$.

Least-squares solution: For a linear system $A\vec{x} = \vec{b}$ with an $n \times m$ matrix A, a vector $\vec{x}^* \in \mathbb{R}^m$ is called a least-squares solution of this system if $||\vec{b} - A\vec{x}^*|| \leq ||\vec{b} - A\vec{x}||$ for $\forall \vec{x} \in \mathbb{R}^m$. The least-squares solutions of the system $A\vec{x} = \vec{b}$ are the exact solutions of the (consistent) system $A^T A\vec{x} = A^T \vec{b}$.

Normal equation: For a system $A\vec{x} = \vec{b}$, the system $A^T A\vec{x} = A^T \vec{b}$ is called the normal equation of $A\vec{x} = \vec{b}$.

Moore-Penrose pseudoinverse: For any $n \times m$ matrix A, there exists a unique matrix A^+ which satisfies the following Penrose equations (i) $AA^+A = A$, (ii) $A^+AA^+ = A^+$, (iii) $(AA^+)^T = AA^+$, (iv) $(A^+A)^T = A^+A$. The matrix A^+ is called the Moore-Penrose pseudoinverse.

Inner product and inner product space: An inner product in a linear space V is a rule that assigns a real scalar $\langle f, g \rangle$ to any pair $f, g \in V$ such that the following properties hold for $\forall f, g, h \in V$, and $\forall c \in \mathbb{R}$: (i) $\langle f, g \rangle = \langle g, f \rangle$. (ii) $\langle f + h, g \rangle = \langle f, g \rangle + \langle h, g \rangle$. (iii) $\langle cf, g \rangle = c\langle f, g \rangle$. (iv) $\langle f, f \rangle > 0$ for all nonzero $f \in V$. A linear space endowed with an inner product is called an inner product space.

Fourier coefficients: Let $f(t)$ be a piecewise continuous function defined on the interval $[-\pi, \pi]$. When $f(t)$ is approximated as $f(t) = \text{proj}_{T_n}(f(t)) = a_0 \frac{1}{\sqrt{2}} + b_1 \sin(t) + c_1 \cos(t) + \cdots + b_n \sin(nt) + c_n \cos(nt)$, the coefficients b_k, c_k, and a_0 are called the Fourier coefficients of $f(t)$.

Several Theorems

Theorem 2.4.8

Let A and B be $n \times n$ matrices such that $BA = I_n$. Then (i) A and B are invertible, (ii) $A^{-1} = B$ and $B^{-1} = A$, and (iii) $AB = I_n$.

Theorem 2.4.9

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, then $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Theorem 3.1.7

For an $n \times m$ matrix A, $\ker(A) = \{\vec{0}\}$ if and only if $\operatorname{rank}(A) = m$.

Theorem 3.1.8, Theorem 3.3.10, and HW3 6)

For an $n \times n$ matrix A, the following statements are equivalent. (i) A is invertible. (ii) $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$ for $\forall \vec{b} \in \mathbb{R}^n$. (iii) $\operatorname{rref}(A) = I_n$. (iv) $\operatorname{rank}(A) = n$. (v) $\operatorname{image}(A) = \mathbb{R}^n$. (vi) $\ker(A) = \{\vec{0}\}$. (vii) The column vectors of A form a basis of $\mathbb{R}^n$. (viii) The column vectors of A span $\mathbb{R}^n$. (ix) The column vectors of A are linearly independent.

Theorem 3.2.9 and HW3 3)

Vectors $\vec{v}_1, \ldots, \vec{v}_m \in \mathbb{R}^n$ are said to be linearly independent if none of $\vec{v}_1, \ldots, \vec{v}_m$ is redundant. The following statements are equivalent. (i) None of $\vec{v}_i$ $(i = 1, \ldots, m)$ is a linear combination of other vectors. (ii) There is only the trivial relation among $\vec{v}_1, \ldots, \vec{v}_m$. (iii) $\ker[\vec{v}_1 \ldots \vec{v}_m] = \{\vec{0}\}$. (iv) $\operatorname{rank}[\vec{v}_1 \ldots \vec{v}_m] = m$.

Theorem 3.2.10 and HW3 4)

The vectors $\vec{v}_1, \ldots, \vec{v}_m$ in a subspace V of $\mathbb{R}^n$ form a basis of V if and only if every vector $v \in V$ can be expressed uniquely as a linear combination $\vec{v} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$.

Theorem 3.3.2

All bases of a subspace V of $\mathbb{R}^n$ consist of the same number of vectors.

Theorem 3.3.5

If we pick up the column vectors of A that correspond to the columns of $\text{rref}(A)$ containing the leading 1's, then they form a basis of the image of A.

Theorem 3.3.6

For any matrix A, $\dim(\text{image}(A)) = \text{rank}(A)$.

Theorem 3.3.7 (Rank-nullity theorem)

For any $n \times m$ matrix A, $\dim(\ker(A)) + \dim(\text{image}(A)) = m$.

Theorem 4.2.3

Any n-dimensional linear space V is isomorphic to $\mathbb{R}^n$.

Theorem 4.3.2

Let B be the matrix of a linear transformation T from an n-dimensional linear space V to V with respect to a basis $\mathcal{B} = (\vec{v}_1, \ldots, \vec{v}_n)$ of V. Then, $B = \left[[T(\vec{v}_1)]_{\mathcal{B}} \ldots [T(\vec{v}_n)]_{\mathcal{B}} \right]$.

Theorem 4.3.4

Consider the change of basis matrix $S_{\mathcal{B}\to\mathcal{A}}$ from a basis $\mathcal{B} = (b_1, \ldots, b_m)$ to another basis $\mathcal{A} = (a_1, \ldots, a_m)$ of a subspace V of $\mathbb{R}^n$. The change of basis matrix $S_{\mathcal{B}\to\mathcal{A}}$ is given by $S_{\mathcal{B}\to\mathcal{A}} = \left[[b_1]_{\mathcal{A}} \ldots [b_n]_{\mathcal{A}} \right]$ and satisfies $\left[\vec{b}_1 \ldots \vec{b}_m \right] = \left[\vec{a}_1 \ldots \vec{a}_m \right] S_{\mathcal{B}\to\mathcal{A}}$.

Theorem 4.3.5

Let $\mathcal{A}$ and $\mathcal{B}$ be bases of a linear space V. Let A and B be the $\mathcal{A}$- and $\mathcal{B}$-matrix of a linear transformation T from V to V. Let $S_{\mathcal{B}\to\mathcal{A}}$ be the change of basis matrix. Then A is similar to B, and $AS_{\mathcal{B}\to\mathcal{A}} = S_{\mathcal{B}\to\mathcal{A}}B$ or $A = S_{\mathcal{B}\to\mathcal{A}}BS_{\mathcal{B}\to\mathcal{A}}^{-1}$ or $B = S_{\mathcal{B}\to\mathcal{A}}^{-1}AS_{\mathcal{B}\to\mathcal{A}}$.

Theorem 5.1.4

For a subspace V of $\mathbb{R}^n$, any vector $\vec{x} \in \mathbb{R}^n$ can be uniquely written as $\vec{x} = \vec{x}_{\parallel} + \vec{x}_{\perp}$, where $\vec{x}_{\parallel}$ is in V and $\vec{x}_{\perp}$ is perpendicular to V.

Theorem 5.1.5, Theorem 5.1.6

If V is a subspace of $\mathbb{R}^n$ with an orthonormal basis $\vec{u}_1, \ldots, \vec{u}_m$, then $\operatorname{proj}_V(\vec{x}) = \vec{x}_{||} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_m \cdot \vec{x})\vec{u}_m$ for all $\vec{x} \in \mathbb{R}^n$. In particular, any $\vec{x} \in \mathbb{R}^n$ is given by $\vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_n \cdot \vec{x})\vec{u}_n$, where $\vec{u}_1, \ldots, \vec{u}_n \in \mathbb{R}^n$ form an orthonormal basis.

Theorem 5.1.8

Consider a subspace V of $\mathbb{R}^n$. (i) $V^{\perp}$ is a subspace of $\mathbb{R}^n$. (ii) $V \cap V^{\perp} = \{\vec{0}\}$. (iii) $\dim(V) + \dim(V^{\perp}) = n$. (iv) $(V^{\perp})^{\perp} = V$.

Theorem 5.1.9 and HW4 7) (Ex. 5.1.12)

For $\vec{x}, \vec{y} \in \mathbb{R}^n$, the equation $||\vec{x} + \vec{y}||^2 = ||\vec{x}||^2 + ||\vec{y}||^2$ holds if and only if $\vec{x}$ and $\vec{y}$ are orthogonal (the Pythagorean theorem). For $\vec{x}$ and $\vec{y}$, the triangle inequality $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$ always holds.

Theorem 5.1.10

Consider a subspace V of $\mathbb{R}^n$ and a vector $\vec{x} \in \mathbb{R}^n$. Then, $||\operatorname{proj}_V(\vec{x})|| \leq ||\vec{x}||$.

Theorem 5.1.11 and HW4 6) (Cauchy-Schwarz inequality)

If $\vec{x}, \vec{y} \in \mathbb{R}^n$, then $|\vec{x} \cdot \vec{y}| \leq ||\vec{x}||\,||\vec{y}||$.

Theorem 5.2.1 (Gram-Schmidt process)

Consider a basis $\vec{v}_1, \ldots, \vec{v}_m$ of a subspace V of $\mathbb{R}^n$. For $j = 2, \ldots, m$, we resolve the vector $\vec{v}_j$ into its components parallel and perpendicular to the span of the preceding vectors $\vec{v}_1, \ldots, \vec{v}_{j-1}$: $\vec{v}_j = \vec{v}_j^{||} + \vec{v}_j^{\perp}$ with respect to $\operatorname{span}(\vec{v}_1, \ldots, \vec{v}_{j-1})$. Then $\vec{u_1} = \vec{v}_1/||\vec{v}_1||$, $\vec{u_2} = \vec{v}_2^{\perp}/||\vec{v}_2^{\perp}||, \ldots, \vec{u_m} = \vec{v}_m^{\perp}/||\vec{v}_m^{\perp}||$ form an orthonormal basis of V. We have $\vec{v}_j^{\perp} = \vec{v}_j - \vec{v}_j^{||} = \vec{v}_j - (\vec{u}_1 \cdot \vec{v}_j)\vec{u}_1 - \cdots - (\vec{u}_{j-1} \cdot \vec{v}_j)\vec{u}_{j-1}$.

Theorem 5.2.2 (QR factorization)

Consider an $n \times m$ matrix M with linearly independent columns $\vec{v}_1, \ldots, \vec{v}_m$. Then there exists an $n \times m$ matrix Q whose columns $\vec{u}_1, \ldots, \vec{u}_m$ are orthonormal and an upper triangular matrix R with positive diagonal entries such that $M = QR$. This representation is unique. Furthermore, $r_{11} = ||\vec{v}_1||$, $r_{jj} = ||\vec{v}_j^{\perp}||$ $(j = 2, \ldots, m)$, and $r_{ij} = \vec{u}_i \cdot \vec{v}_j$ $(i < j)$.

Theorem 5.3.3, Theorem 5.3.7, and Theorem 5.3.8

An $n \times n$ matrix A is an orthogonal matrix if $||A\vec{x}|| = ||\vec{x}||$ for $\forall \vec{x} \in \mathbb{R}^n$. The following statements are equivalent: (i) The columns of A from an orthonormal basis of $\mathbb{R}^n$. (ii) $A^T A = I_n$. (iii) $A^{-1} = A^T$.

HW5 5) (Ex. 5.3.28)

An orthogonal transformation L from $\mathbb{R}^n$ to $\mathbb{R}^n$ preserves the dot product: $\vec{v} \cdot \vec{w} = L(\vec{v}) \cdot L(\vec{w})$, for all $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$.

HW5 10) (Ex. 5.4.16)

$\text{rank}(A) = \text{rank}(A^T)$.

Theorem 5.4.1

For any matrix A, we have $[\text{image}(A)]^\perp = \ker(A^T)$.

Theorem 5.4.2

If A is an $n \times m$ matrix, then $\ker(A) = \ker(A^T A)$. If A is an $n \times m$ matrix with $\ker(A) = \{\vec{0}\}$, then $A^T A$ is invertible.

Theorem 5.4.6

If $\ker(A) = \{\vec{0}\}$, then the linear system $A\vec{x} = \vec{b}$ has the unique least-squares solution $\vec{x}^* = (A^T A)^{-1} A^T \vec{b}$.

(Moore-Penrose pseudoinverse)

For any $n \times m$ matrix A, there exists a unique matrix A^+ which satisfies the following Penrose equations (i) $AA^+A = A$, (ii) $A^+AA^+ = A^+$, (iii) $(AA^+)^T = AA^+$, (iv) $(A^+A)^T = A^+A$. The matrix A^+ is called the Moore-Penrose pseudoinverse. If $n = m = \text{rank}(A)$, then $A^+ = A^{-1}$. If $\text{rank}(A) = m$, then $A^+ = (A^T A)^{-1} A^T$.

Some Proofs

First go back to the previous section and think how to prove the theorem. Then, read the proof of the theorem below.

Proof of Theorem 2.4.8 (i), (ii): Consider $\vec{x}$ such that $A\vec{x} = \vec{0}$. We have $I_n\vec{x} = BA\vec{x} = B\vec{0} = \vec{0}$. Thus $\vec{x} = \vec{0}$. Since $\vec{x}$ is uniquely determined in $A\vec{x} = \vec{0}$, A is invertible (Theorem 3.1.8). Let us operate A^{-1} from right: $BAA^{-1} = I_nA^{-1}$. Hence, $B = A^{-1}$. We obtain $B^{-1} = (A^{-1})^{-1} = A$.

(iii): $AB = AA^{-1} = I_n$.

Proof of Theorem 3.2.10 ($\Longrightarrow$) Suppose the vector $\vec{v}$ is expressed in two ways: $\vec{v} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$ and $\vec{v} = c_1'\vec{v}_1 + \cdots + c_m'\vec{v}_m$. By subtraction, we obtain $(c_1 - c_1')\vec{v}_1 + \cdots + (c_m - c_m')\vec{v}_m = \vec{0}$. However, since $\vec{v}_1, \ldots, \vec{v}_m$ are linearly independent, we have $c_1 = c_1', \ldots, c_m = c_m'$.

($\Longleftarrow$) Consider $\vec{v} = \vec{0}$. Obviously $\vec{0}$ is expressed with $c_1 = \cdots = c_m = 0$. However, it is uniquely expressed, $c_1\vec{v}_1 + \cdots + c_m\vec{v}_m = \vec{0}$ has only the solution $c_1 = \cdots = c_m = 0$. This means $\vec{v}_1, \ldots, \vec{v}_m$ are linearly independent. They span V. Therefore they form a basis.

Proof of Theorem 5.1.4 Consider an orthonormal basis $\vec{u}_1, \ldots, \vec{u}_m \in V$ ($\dim(V) = m \leq n$). We decompose $\vec{x}$ as $\vec{x} = \vec{x}_{\parallel} + \vec{x}_{\perp}$, where $\vec{x}_{\parallel} \in V$. Then we can write $\vec{x}_{\parallel} = c_1\vec{u}_1 + \cdots + c_m\vec{u}_m$. We want to choose $\vec{x}_{\parallel}$ so that $\vec{x}_{\perp} = \vec{x} - \vec{x}_{\parallel}$ is perpendicular to any vector $\vec{v} \in V$. We write $\vec{v} = k_1\vec{u}_1 + \cdots + k_m\vec{u}_m$. Consider $\vec{u}_i \cdot \vec{x}_{\perp} = \vec{u}_i \cdot \vec{x} - c_i$. Therefore, only when $c_i = \vec{u}_i \cdot \vec{x}$ ($i = 1, \ldots, m$), we have $\vec{v} \cdot \vec{x}_{\perp} = 0$. The vector $\vec{x}$ is uniquely decomposed as $\vec{x} = \vec{x}_{\parallel} + \vec{x}_{\perp}$, where $\vec{x}_{\parallel} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_m \cdot \vec{x})\vec{u}_m$.

Proof of Theorem 5.1.11 Let us consider the function $q(t)$ defined by $q(t) = ||\vec{x} + t\vec{y}||^2 \geq 0$. By definition, we have

$$q(t) = (\vec{x} + t\vec{y}) \cdot (\vec{x} + t\vec{y}) = ||\vec{x}||^2 + 2t\vec{x} \cdot \vec{y} + t^2||\vec{y}||^2.$$

Suppose $||\vec{y}|| \neq 0$. We set $t = -\vec{x} \cdot \vec{y}/||\vec{y}||^2$. We obtain $0 \leq q(t) = ||\vec{x}||^2 - (\vec{x} \cdot \vec{y})^2/||\vec{y}||^2$. Therefore, we obtain $|\vec{x} \cdot \vec{y}| \leq ||\vec{x}||\,||\vec{y}||$. If $\vec{x}$ and $\vec{y}$ are parallel, the equality holds because then we can write $\vec{y} = k\vec{x}$ with scalar k. If $\vec{y} = \vec{0}$ and $||\vec{y}|| = 0$, we have $|\vec{x} \cdot \vec{y}| = 0$ and $||\vec{x}||\,||\vec{y}|| = 0$. The case of $||\vec{y}|| = 0$ is included in the above inequality.

Alternative solution Note that $\vec{x} \cdot \vec{y} = ||\vec{x}||\,||\vec{y}|| \cos\theta$. Since $-1 \leq \cos\theta \leq 1$, we have $|\vec{x} \cdot \vec{y}| \leq ||\vec{x}||\,||\vec{y}||$. The equality holds when $\vec{x}$ and $\vec{y}$ are parallel.

Proof of HW5 5) (Ex. 5.3.28) The transformation L is done by an orthogonal matrix $Q = [\vec{u}_1 \ldots \vec{u}_n]$ (Theorem 5.3.3). We have $L(\vec{v}) \cdot L(\vec{w}) = (Q\vec{v}) \cdot (Q\vec{w}) = (v_1\vec{u}_1 + \cdots + v_n\vec{u}_n) \cdot (w_1\vec{u}_1 + \cdots + w_n\vec{u}_n) = v_1w_1 + \cdots + v_nw_n = \vec{v} \cdot \vec{w}$.

Proof of HW5 10) (Ex. 5.4.16) Note that $\dim[\text{image}(A)] = \text{rank}(A)$ (Theorem 3.3.6), $\dim[\ker(A)] + \dim[\text{image}(A)] = m$ (Theorem 3.3.7), and $\dim[\text{image}(A)] + \dim[(\text{image}(A))^{\perp}] = n$ (Theorem 5.1.8). Therefore,

$$\begin{aligned} \text{rank}(A) &= \dim[\text{image}(A)] \\ &= n - \dim[(\text{image}(A))^{\perp}] \\ &= n - \dim[\ker(A^T)] \\ &= n - \big(n - \dim[\text{image}(A^T)]\big) \\ &= \dim[\text{image}(A^T)] \\ &= \text{rank}(A^T). \end{aligned}$$

Proof of Theorem 5.4.1 $V = \text{image}(A)$ is a subspace of $\mathbb{R}^n$. $V^{\perp} = \{\vec{x} : \vec{v}_i \cdot \vec{x} = 0,\, i = 1, \ldots, m\} = \{\vec{x} : \begin{bmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_m^T \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}\} = \{\vec{x} : A^T\vec{x} = \vec{0}\} = \ker(A^T)$.

More Problems

Go over homework problems. Here are more problems if you need. Solutions can be found in the textbook.

Chapter 1

1.1.39, 1.2.3, 1.2.5, 1.2.7, 1.2.9, 1.2.11, 1.2.45, 1.2.49, 1.3.47

Chapter 2

2.1.13, 2.1.37, 2.2.7, 2.2.9, 2.2.13, 2.2.19, 2.2.27, 2.2.33, 2.2.37, 2.2.43, 2.3.29, 2.4.77, 2.4.89

Chapter 3

3.1.39, 3.1.47, 3.1.51, 3.2.35, 3.2.49, 3.3.21, 3.3.23, 3.3.25, 3.3.27, 3.3.29, 3.3.71, 3.3.83, 3.4.25, 3.4.27, 3.4.29, 3.4.47, 3.4.55

Chapter 4

4.1.23, 4.1.25, 4.2.67, 4.2.69, 4.3.1, 4.3.3, 4.3.5, 4.3.7, 4.3.9, 4.3.11, 4.3.13, 4.3.61, 4.3.63

Chapter 5

5.1.21, 5.1.27, 5.2.1, 5.2.3, 5.2.5, 5.2.7, 5.2.9, 5.2.11, 5.2.13, 5.3.33, 5.3.45, 5.3.55, 5.4.19, 5.4.21, 5.5.15, 5.5.23