Linear Algebra: Matrices and Geometry

1 Duality

Definition (Dual Space). If V is a vector space over a field $\mathbb{F}$, the vector space $\mathcal{L}(V, \mathbb{F})$ of all linear forms on V is called the *dual space* of V, denoted V'.

Let $\phi_1, \phi_2 \in V'$, $c \in \mathbb{F}$, and $v \in V$, then $(\phi_1 + \phi_2)(v) = \phi_1(v) + \phi_2(v)$ and $(c\phi_1)(v) = c(\phi_1(v))$.

Proposition (Basis of a Dual Space). Assume V is a finite-dimensional vector space over a field $\mathbb{F}$, with basis $B = \{v_1, v_2, \ldots, v_n\}$. Note that for any $v \in V$,

$$v = \sum_{i=1}^{n} a_i v_i, \quad \text{where } a_i \in \mathbb{F}.$$

Consider a linear form, $\phi \in V'$, then $\phi(v) = a_1\phi(v_1) + a_2\phi(v_2) + \cdots + a_n\phi(v_n)$. Therefore, suppose we define $B' = \{f_1, \ldots, f_n\}$ such that $f_i(v_j) = \delta_{ij}$ for all $i, j = 1, \ldots, n$, where δ_{ij} denotes the Kronecker delta function. Then, B' is a basis for V'.

Proof. First, consider $f_i(v)$ where v is an arbitrary vector in V. Then,

$$\begin{aligned} f_i(v) &= f(a_1v_1 + a_2v_2 + \cdots + a_nv_n) \\ &= a_1f_i(v_1) + \cdots + a_nf_i(v_n) \\ &= a_1\delta_{i1} + \cdots + a_n\delta_{in} \\ &= a_i. \end{aligned}$$

So we see that f_i outputs the i^{th} coordinate of v. For any $\phi \in V'$, consider $\phi(v_1)f_1 + \cdots + \phi(v_n)f_n$:

$$\begin{aligned} (\phi(v_1)f_1 + \cdots + \phi(v_n)f_n)(v) &= (\phi(v_1)f_1)(v) + \cdots + (\phi(v_n)f_n)(v) \\ &= \phi(v_1)(f_1(v)) + \cdots + \phi(v_n)(f_n(v)) \\ &= \phi(v_1)a_1 + \cdots + \phi(v_n)a_n \\ &= \phi(v). \end{aligned}$$

Thus, B' is clearly generating for V'. Now we show, f_i are independent. Let v_i be the i^{th} basis vector of V and $\mathbf{0} : V \to \mathbb{F}$ such that $v \mapsto 0$. Suppose $c_1f_1 + \cdots + c_nf_n = \mathbf{0}$, then

$$\begin{aligned} (c_1f_1 + \cdots + c_nf_n)(v_i) &= \mathbf{0}(v) \\ c_1(f_1(v_i)) + \cdots + c_n(f_n(v_i)) &= 0 \\ c_1 \cdot 0 + \cdots + c_i \cdot 1 + \cdots + c_n \cdot 0 &= 0. \end{aligned}$$

Thus for any $i = 1, \ldots, n$ we can show c_i must be 0. ∎

Example. Let $V = \mathbb{R}^2$, with basis $\{\mathbf{e}_1, \mathbf{e}_2\}$. Then, $B' = \{f_1, f_2\}$, such that for any $(x, y) \in \mathbb{R}^2$, $f_1((x, y)) = x$ and $f_2((x, y)) = y$, is a basis for the dual space of V.

Corollary S. ince the basis of V and V' have the number of elements, $\dim(V) = \dim(V')$ and $V \cong V'$ via the mapping $v \mapsto \phi$.

Definition (Transpose of a Linear Mapping). Given any linear mapping $T : V \to W$, we can define its *transpose* $T' : W' \to V'$ by the mapping $T'g = gT$, for any $g \in \mathcal{L}(W, \mathbb{F})$.

As an exercise, prove that the above mapping is indeed linear.

Theorem. $\mathrm{Ker}(T') = \{g \in W' | g = 0 \text{ on } T(V)\}$

Proof. For $g \in W'$, $g = 0$ on $\mathrm{Im}(T) \Leftrightarrow g(T(V)) = 0 \Leftrightarrow (gT)(V) = 0 \Leftrightarrow (T'g)(V) = 0 \Leftrightarrow T'g = 0 \Leftrightarrow g \in \mathrm{Ker}(T')$. ∎

Definition (Annihilator). Let V be a vector space with subspace M. The *annihilator* of M in V', denoted M° is the set of all linear forms f on V such that $f(x) = 0$ for all $x \in M$.

Note that for any linear map $T : V \to W$ the $\text{Ker}(T') = (T(V))^\circ$.

As an exercise, with notations as in definition 1.3, prove that $\dim(M) + \dim(M^\circ) = \dim(V)$. {Hint: Use the First Isomorphism theorem and let suppose for some linear mapping T, the $\text{Ker}(T') = M^\circ$ }

Definition (Bidual Space). Let V be a vector space. The *bidual* of V, denoted $(V')'$ is the set of linear forms from V' onto $\mathbb{F}$.

Notice that there is a natural injection, $V \xmapsto{\Psi} V''$. Specifically, $\Psi(x) : V' \to \mathbb{F}$, for some $x \in V$ where $(\Psi(x))(\alpha) = \alpha(x)$ and $\alpha \in V'$. Clearly, $\text{Ker}(\Psi) = \{0 \in V\}$. So Ψ is injective. For a fixed $x \in V$, $\Psi(x)$ is linear:

Proof. Let $\alpha_1, \alpha_2 \in V$. Then, $(\Psi(x))(\alpha_1 + \alpha_2) = (\alpha_1 + \alpha_2)(x) = \alpha_1(x) + \alpha_2(x) = (\Psi(x))(\alpha_1) + (\Psi(x))(\alpha_2)$.
Let $\alpha \in V$ and $c \in \mathbb{F}$. Then, $(\Psi(x))(c\alpha) = (c\alpha)(x) = c(\alpha(x)) = c(\Psi(x)(\alpha))$. ■

Therefore $V \cong V' \cong V''$.

2 Matrices and Linear Transformations

Definition 2.1: Matrix of a Linear Transformation

Suppose $T : V \to W$ is a linear transformation. Choose bases of $x_1, x_2, \ldots, x_n$ and $y_1, \ldots, y_m$ of V, W, respectively. Then for each $j = 1, \ldots, n$ express Tx_j as a linear combination of y_i:

$$Tx_j = \sum_{i=1}^{m} a_{ij} y_i.$$

The $m \times n$ matrix, (a_{ij}), is called the matrix of T relative to the bases of V and W. Note that this matrix isn't necessarily unique since it's dependent on the choice of bases for V, W.

Example. The $n \times n$ (identity) matrix of the identity mapping, $I : V \to V$, is given by:

$$(\delta_{ij}) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Definition (Column/Row Rank). Let A be an $m \times n$ matrix over $\mathbb{F}$. The *column space* of A is the linear subspace of $\mathbb{F}^m$ generated by the column vectors of A. It's dimension is called the *column rank*. *Row space* and *row rank* are defined analogously.

Theorem. If $T : V \to \mathbb{F}^m$ a linear mapping and A its matrix relative to the bases for V and $\mathbb{F}^m$, then $T(V)$ is the linear span of the column vectors of A.

Theorem. If $T : V \to W$ a linear mapping and A its matrix relative to the bases for V and W, then the rank of T is equal to the column rank of A.

2.1 Matrix Multiplication

Theorem. Let U, V, W be vector spaces over $\mathbb{F}$ and let $T : U \to V$ and $S : V \to W$ be linear, with ST the composite mapping. Choose bases for U, V, W respectively and let A be the matrix of S and B that of T, relative to their bases. Then, the matrix of ST is AB.

Proof. Let $p = \dim(U)$, $n = \dim(V)$, $m = \dim(W)$, and choose bases:

$$\begin{gathered} z_1, z_2, \ldots, z_p \text{ of } U \\ x_1, x_2, \ldots, x_n \text{ of } V \\ y_1, y_2, \ldots, y_m \text{ of } W. \end{gathered}$$

Therefore,

$$Sx_j = \sum_{i=1}^{m} a_{ij} y_i \quad \text{for } j = 1, \ldots, n.$$

$$Tz_k = \sum_{j=1}^{n} b_{jk} x_j \quad \text{for } k = 1, \ldots, p.$$

$$STz_k = \sum_{i=1}^{m} c_{ik} y_i \quad \text{for } k = 1, \ldots, p.$$

and the matrices of S, T, ST are

$$A = (a_{ij}) \qquad B = (b_{jk}) \qquad C = (C_{ik}),$$

respectively. Note A is $m \times n$ and B is $n \times p$ so multiplication is defined and AB is $m \times p$, as is C. For every $k = 1, \ldots, p$;

$$\begin{aligned}
\sum_{i=1}^{m} c_{ik} y_i &= (ST) z_k = S(Tz_k) \\
&= S(\sum_{j=1}^{n} b_{jk} x_j) = \sum_{j=1}^{n} b_{jk} S(x_j) \\
&= \sum_{j=1}^{n} b_{jk} \Big(\sum_{i=1}^{m} a_{ij} y_i \Big) \\
&= \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij} b_{jk} \Big) y_i.
\end{aligned}$$

Hence, $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$ for all i, k; therefore, $C = AB$. ■

2.2 Transpose

Definition (Tranpose of a Matrix). Let $A = (a_{ij})_{1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n}$ be an $m \times n$ matrix. The transpose of A denoted A^{T} is the $n \times m$ matrix (b_{ji}) for which $b_{ji} = a_{ij}$.

Therefore, the row space of A is equal to the column space of A^{T} and vice-versa.

Theorem. Let V, W be finite-dimensional vector spaces, $T : V \to W$ a linear mapping, and $T' : W' \to V'$ its transpose. Choose bases for V, W and construct bases for V', W' dual to them. Then, relative to these bases, the matrix of T' is the transpose of the matrix of T.

Corollary 2.4. For an arbitrary matrix A, row rank(A) = column rank(A).

2.3 Rank of a Matrix

Rank Preserving Operations:

I. Permuting rows or columns, denoted $r_i \leftrightarrow r_j$.

II. Adding a scalar multiple of one row (column) to another row (column), denoted $r_i \leftarrow r_i + cr_j$.

III. Multiplying a row (column) by a non-zero scalar, denoted $r_i \leftarrow cr_i$.

Theorem. If A is a non-zero matrix, then by performing a finite number of the above operations, A can be brought into the form:

$$\left(\begin{array}{c|c} I & O \\ \hline O & O \end{array} \right)$$

where I is the $r \times r$ identity matrix, and the Os denote zero matrices of the appropriate size. Thus, rank$(A) = r$. This is called the reduced row echelon form of the matrix A, and we denote it rref(A).

2.4 Change of Basis and Matrix Invertibility

Proposition (Change of Basis). Suppose $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ are two bases of an n-dimensional vector space V. Then, the vectors in each basis can be expressed in terms of the other:

$$y_j = \sum_{i=1}^{n} a_{ij} x_i, \quad x_j = \sum_{i=1}^{n} b_{ij} y_i \quad (j = 1, \ldots, n).$$

If $A = (a_{ij})$ and $B = (b_{ij})$ are the coefficient matrices, then $AB = BA = I$.

We call A and B the change-of-basis matrices. For instance, multiplication by A transforms some y_j into its equivalent representation in the x basis.

Definition (Matrix Inverse). An $n \times n$ matrix A is invertible if and only if there exists an $n \times n$ matrix B such that $AB = BA = I$.

Such a matrix is unique, since if $AC = I$ then $C = IC = (BA)C = B(AC) = BI = B$. Thus we call B the inverse of matrix A, denoted A^{-1}.

Theorem. Let A be the matrix of the linear mapping $T : V \to V$ relative to some basis for V. Then, A is invertible if and only if T is bijective.

Theorem. Let V be an n-dimensional vector space and $T : W \to V$ a linear mapping. Let $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ be bases of V and A, B be the matrices of T with respect to the x_i, y_j bases, respectively. If C is the change of basis matrix that expresses the y_j in terms of the x_i, then $B = C^{-1}AC$.

Definition (Similar Matrices). Matrices A, B are *similar* is there exists an invertible matrix C such that $B = C^{-1}AC$.

3 Inner Product Spaces

Definition (Inner Product Space). An *inner product space* is a real vector space $\mathbb{E}$ such that, for every pair of vectors $x, y \in \mathbb{E}$, there is determined a real number $(x|y)$ called the inner product of x and y, subject to the following axioms:

1. $(x|x) > 0$ for all $x \neq \theta$.
2. $(x|y) = (y|x)$ for all x, y.
3. $(x + y|z) = (x|z) + (y|z)$ for any x, y, z.
4. For any scalar c, $(cx|y) = c(x|y)$ for any x, y

Example. The canonical inner product on $\mathbb{R}^n$ is defined by

$$(x|y) = \sum_{i=1}^{n} a_i b_i$$

for all $x = (a_1, \ldots, a_n)$ and $y = (b_1, \ldots, b_n)$.

Let $[a, b]$ be a closed interval on $\mathbb{R}$ with $a < b$ and E the vector space of all continuous functions $x : [a, b] \to \mathbb{R}$, then the with the inner product defined as

$$(x|y) = \int_a^b x(t)y(t)dt.$$

Theorem. In an inner product space, the following identities hold:

1. $(x|y + z) = (x|y) + (x|z)$
2. $(x|cy) = c(x|y)$
3. $(x|\theta) = (\theta|y) = 0$
4. $||cx|| = |c| \cdot ||x||$
5. $(x|y) = \frac{1}{4}[||x + y||^2 - ||x - y||^2]$

6. $||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2$

Theorem. In an inner product space,

$$|(x|y)| \leqslant ||x||||y||$$

for all vectors x, y; with equality if and only if x and y are linearly dependent.

Proof. If $x = \theta$ or $y = \theta$ the inequality holds trivially. Assume x, y are nonzero, let

$$u = ||x||^{-1}x \qquad v = ||y||^{-1}y,$$

thus u, v are unit vectors and

$$(x|y) = ((||x||u) \mid (||y||v)) = ||x||||y||(u|v).$$

So it suffices to show $|(u|v)| \leqslant 1$. By Theorem 3.1, (5), (6) we have

$$\begin{aligned} |(u|v)| &\leqslant \frac{1}{4}[||u+v||^2 + ||u-v||^2] \\ &= \frac{1}{4}[2||u||^2 + 2||v||^2] = 1. \end{aligned}$$

■

Theorem. In an inner product space,

$$||x+y|| \leq ||x|| + ||y||.$$

Definition (Orthogonal). Two vectors are orthogonal or perpendicular, written $x \perp y$, if $(x|y) = 0$. Subsets A and B of an inner product space $\mathbb{E}$ are orthogonal if $x \perp y$ for all $x \in A$ and $y \in B$.

Theorem. If $x_1, x_2, \ldots, x_n$ are pairwise orthogonal vectors in an inner product space, then they are linearly independent.

Proof. Assuming $c_1x_1 + c_2x_2 + \cdots + c_nx_n = \theta$ we have

$$0 = (\theta|x_j) = \left(\sum_{i=1}^{n}(c_ix_i|x_j)\right) = \sum_{i=1}^{n} c_i(x_i|x_j) = c_j||x_j||^2.$$

■

Definition (Annihilator). The set of all vectors $x \in \mathbb{E}$ such that $x \perp A$ is denoted $A^\perp$ and is called the annihilator of A in $\mathbb{E}$.

3.1 Duality in Inner Product Spaces

Definition (Linear Form in the IPS). If $\mathbb{E}$ is an inner product space and $y \in \mathbb{E}$, then a natural linear form to define is

$$y'(x) = (x|y)$$

for all $x \in \mathbb{E}$.

Theorem. If $\mathbb{E}$ is an inner product space, then the mapping $\mathbb{E} \to \mathbb{E}'$ defined by $y \mapsto y'$ is linear and injective.

Proof. Since $(x|y+z) = (x|y) + (x|z)$, $(y+z)' = y' + z'$ and since $(x|cy) = c(x|y)$ we have $(cy)' = cy'$. If $y' = 0$ then $0 = y'(y) = (y|y)$ which implies that $y = \theta$. Hence, the mapping has kernel θ. ■

Theorem. If $\mathbb{E}$ is a Euclidean space, then the mapping $y \mapsto y'$ is a vector space isomorphism $\mathbb{E} \to \mathbb{E}'$.

Proof. Since $\dim \mathbb{E}' = \dim \mathbb{E}$, the linear mapping $y \mapsto y'$ which is injective, is necessarily bijective. ■

Definition (Canonical Isomorphism). The linear bijection $J : \mathbb{E} \to \mathbb{E}'$ defined by $Jy = y'$ is the canonical isomorphism of the Euclidean space $\mathbb{E}$ onto its dual.

Theorem. If M is a finite-dimensional linear subspace of an inner product space $\mathbb{E}$ then $E = M \oplus M^\perp$.

Proof. Since $M, M^\perp$ are linear subspace they both contain θ, thus $\theta \subset M \cap M^\perp$. Conversely, if $x \in\subset M \cap M^\perp$ then $x \perp x$, so $x = \theta$. Thus, $\subset M \cap M^\perp = \theta$.

To show $\mathbb{E} = M + M^\perp$, we must show that given any $x \in \mathbb{E}$, we can find suitable $y \in M$ and $z \in M^\perp$ s.t. $x = y + z$. Let x' be a linear form on $\mathbb{E}$ as in Def. 3.4. Let f be the restriction of x' to M. Then, f is a linear form on Euclidean space M, so there exists a vector $y \in M$ s.t. $f(w) = (w|y)$ for all $w \in M$. Hence $(w|x) = (w|y)$ for all $w \in M$. Then, $(w|x - y) = 0$ for all $w \in M$ so $z = x - y \in M^\perp$. ∎

Corollary 3.7. $\dim M^\perp = \dim \mathbb{E} - \dim M$.

Definition (Orthogonal Complement). If $\mathbb{E}$ is a Euclidean space with linear subspace M, then $M^\perp$ is the *orthogonal complement* of M in $\mathbb{E}$.

Definition (Orthonormal). Vectors $x_1, x_2, \ldots, x_n$ in an inner product space are *orthonormal* if they are pairwise orthogonal unit vectors.

Theorem. Every Euclidean space $\neq \{\theta\}$ has an orthonormal basis.

Corollary 3.8. If $\dim \mathbb{E} = n$, then there exists a bijective linear mapping $T : \mathbb{E} \to \mathbb{E}^n$ such that $(Tx|Ty) = (x|y)$ for all $x, y \in \mathbb{E}$. Specifically, $T = S^{-1}$, where $x_1, x_2, \ldots, x_n$ is an orthonormal basis of $\mathbb{E}$ and $S(a_1, a_2, \ldots, a_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$.

Definition (Isomorphism). Two inner product spaces $\mathbb{E}, \mathbb{F}$ are isomorphic if there exists a bijective linear mapping $T : \mathbb{E} \to \mathbb{F}$ such that $(Tx|Ty) = (x|y)$ for all $x, y \in \mathbb{E}$.

3.2 Adjoint

Definition (Adjoint). The *adjoint* of a linear map $T : \mathbb{E} \to \mathbb{F}$ is $T^* : \mathbb{F} \to \mathbb{E}$ defined by $T^* = J_{\mathbb{E}}^{-1} T' J_{\mathbb{F}}$.

Theorem. Let $T : \mathbb{E} \to \mathbb{F}$, then

1. T^* is linear
2. $(Tx|y) = (x|T^*y)$ and $(T^*y|x) = (y|Tx)$ for all $x \in \mathbb{E}$, $y \in \mathbb{F}$.
3. If $S : \mathbb{F} \to \mathbb{E}$ is a mapping such that $(Tx|y) = (x|Sy)$ for all $x \in \mathbb{E}$, $y \in \mathbb{F}$ then necessarily $S = T^*$.
4. $(T^*)^* = T$
5. $(S + T)^* = S^* + T^*$ and $(cT)^* = cT^*$
6. $(ST)^* = T^*S^*$
7. $I^* = I$ and $0^* = 0$
8. $TT^* = 0$ or $T^*T = 0$ implies $T = 0$.

Corollary 3.9. If $T : \mathbb{E} \to \mathbb{F}$ is a linear bijection, then $T^* : \mathbb{F} \to \mathbb{E}$ is bijective and $(T^*)^{-1} = (T^{-1})^*$.

Proof. Taking adjoints in $T^{-1}T = I$ and $TT^{-1} = I$ we have $T^*(T^{-1})^* = I$ and $(T^{-1})^*T^* = I$. ∎

Theorem. Let $T : \mathbb{E} \to \mathbb{F}$ be linear. Choose orthonormal bases of $\mathbb{E}$ and $\mathbb{F}$ and let A be the matrix of T relative to these bases. Then the matrix of $T^* : \mathbb{F} \to \mathbb{E}$ relative to these bases is the transpose A^T of A.

3.3 Orthogonal Mappings

Definition (Orthogonal Mapping). Let E be a Euclidean space and $T \in \mathcal{L}(E)$. A linear mapping $T \in \mathcal{L}(E)$ is said to be orthogonal if $(Tx|Ty) = (x|y)$ for all $x, y \in E$.

Let E be a Euclidean space and $T \in \mathcal{L}(E)$, then the following are equivalent:

- T is orthogonal
- $||Tx|| = ||x||$ for all $x \in E$
- $T^*T = I$

- $TT^* = I$.

If A is the matrix of T relative to an orthonormal basis of E, then T is orthogonal if and only if $A^\mathsf{T} A = I$. Hence, we say an $n \times n$ matrix A is *orthogonal* if $A^\mathsf{T} A = I$.

For a real $n \times n$ matrix A the following are equivalent: (a) A is orthogonal; (b) A is invertible and $A^{-1} = A^\mathsf{T}$; (c) A^T is orthogonal.

4 Similarity

Matrices A and B are similar if and only if there exists a linear mapping $T : V \to V$ such that A and B are the matrices of T relative to two bases of V. Furthermore, if $A \sim B$, then $|A| = |B|$ and $\mathrm{tr}(A) = \mathrm{tr}(B)$.

Proposition (Matrix Commutativity). If $C^{-1}AC = A$ for all invertible matrices C, then $A = aI$ for some $a \in \mathbb{F}$.

Theorem. If the field F is algebraically closed, then every matrix $A \in \mathrm{Mat}_n(F)$ is similar to a triangular matrix B.

Proof. Let V be an n-dimensional vector space over F, choose a basis of B and let $T \in \mathcal{L}(V)$ be the linear mapping whose matrix relative to the chosen basis is A. For $n = 1$, there is nothing to prove. Since F is algebraically closed, p_T has a root $c_1 \in F$. So c_1 is an eigenvalue of T. Let x_1 be an eigenvector corresponding to c_1 and $M = Fc_1$ the 1-dimensional subspaces spanned by x_1. The quotient space V/M is $(n-1)$-dimensional; let $Q : V \to V/M$ be the quotient mapping.

Since $T(M) \subset M$ we can define $S : V/M \to V/M$ as follows. Let $u \in V/M$, say $u = Qx$, and define $Su = Q(Tx)$. Note S is well-defined and linear. By the inductive hypothesis, V/M has a basis $u_2, \ldots, u_n$ such that for each $j \geq 2$, Su_j is a linear combination of $u_2, \ldots, u_j$, say

$$Su_j = b_{2j}u_2 + \cdots b_{jj}u_j.$$

Choose $x_j \in V$ with $Qx_j = u_j$, $j = 2, \ldots, n$. Then $Tx_j = b_{ij}x_1 + \cdots + b_{jj}x_j$ for $j = 2, \ldots, n$. Writing $b_{11} = c_1$, we have $Tx_1 = b_{11}x_1$. Finally, $x_1, \ldots, x_n$ generate V and are thus a basis for V. ■