LA2 3

KYB

Thrn, it's a Fact

mathrnfact@gmail.com

September 25, 2020

Overview

Project : Pseudoinverse

Ex6.6.13, the minimum-norm solution to $Ax = y$

Let $A \in \mathbb{R}^{m\times n}$ and $y \in \mathbb{R}^m$ be given. Suppose A is singular and $y \in \mathrm{col}(A)$. Then $Ax = y$ has infinitely many solutions and the solution set is $\hat{x} + \mathcal{N}(A)$, where $\hat{x}$ is any one solution.

(a) Prove that there is a unique solution $\overline{x}$ to $Ax = y$ such that $\overline{x} \in \mathrm{col}(A^T)$.

(b) Prve that if $x \in \mathbb{R}^n$ is a solution to $Ax = y$ and $x \neq \overline{x}$, then $\|\overline{x}\|_2 < \|x\|_2$.

Ex6.6.14, the minimum-norm least-squares solution to $Ax = y$

The set of all least squares solutions to $Ax = y$ is $\hat{x} + \mathcal{N}(A)$, where $\hat{x} \in \mathbb{R}^n$ is any one least-squares solution. Prove that $\overline{x}$ has the smallest Euclidean norm of any element of $\hat{x} + \mathcal{N}(A)$.

Remark

Suppose $y \in \operatorname{col}(A)$. Let $\hat{x}$ be a least-square solution to $Ax = y$ and $\tilde{x}$ be a solution to $Ax = y$. Then for any $x \in \mathbb{R}^n$,

$$\|A\hat{x} - y\| \leq \|Ax - y\|.$$

In particular,

$$\|A\hat{x} - y\| \leq \|A\tilde{x} - y\| = 0.$$

Hecne $\hat{x}$ is a solution to $Ax = y$.
So the least-square solution is a generalization of the soltuion to $Ax = y$.

Ex6.6.15, the pseudoinverse of A

Define $S : \mathbb{R}^m \to \mathbb{R}^n$ as follows: $\overline{x} = S(y)$ is the minimum-norm least-squares solution to $Ax = y$.

(a) Prove that S is a linear operator. It follows that there is a matrix $A^\dagger \in \mathbb{R}^{n\times m}$ such that $S(y) = A^\dagger y$.

(b) Find formulas for $A^\dagger$ in each of the following cases:

   i. $A \in \mathbb{R}^{n\times n}$ is non singular.
   ii. $A \in \mathbb{R}^{m\times n}$, $m > n$, has full rank.
   iii. $A \in \mathbb{R}^{m\times n}$, $m < n$, has rank m.

Ex6.6.16

Find $\mathcal{N}(A^\dagger)$.

Ex6.6.17

Prove that $\operatorname{col}(A^{\dagger}) = \operatorname{col}(A^{T})$.

Ex6.6.18

Prove that $AA^\dagger$ is the matrix defining the orthogonal projection onto $\mathrm{col}(A)$.

Ex6.6.19

Prove that $A^\dagger A$ is the matrix defining the orthogonal projection onto $\mathrm{col}(A^T)$.

Recall, Ex6.2.6

If $A, B \in \mathbb{R}^{m \times n}$ and

$$y \cdot Ax = y \cdot Bx \text{ for } x \in \mathbb{R}^n, y \in \mathbb{R}^m,$$

then $A = B$.

Proof.

Since $A_{ij} = e_i \cdot (Ae_j)$,

$$A_{ij} = e_i \cdot (Ae_j) = Be_i \cdot (Ae_j) = B_{ij}.$$

$\square$

Ex6.6.20

Prove that the following equations hold for all $A \in \mathbb{R}^{m \times n}$:

(a) $AA^{\dagger}A = A$;

(b) $A^{\dagger}AA^{\dagger} = A^{\dagger}$;

(c) $A^{\dagger}A = (A^{\dagger}A)^T$;

(d) $AA^{\dagger} = (AA^{\dagger})^T$.

Recall

In Ex6.4.16 and Ex6.4.17, we see that a linear map $P : V \to V$ is an orthogonal projection operator if and only if $P^2 = P$ and $P^* = P$.

Ex6.6.21

Prove that the unique matrix $B \in \mathbb{R}^{n \times m}$ satisfying

$$ABA = A, BAB = B, BA = (BA)^T, AB = (AB)^T$$

is $B = A^\dagger$. Hence, $(A^\dagger)^\dagger = A$.

Properties of the pseudoinverse

- If A is invertible, $A^\dagger = A^{-1}$.
- $(A^\dagger)^\dagger = A$.
- $(A^T)^\dagger = (A^\dagger)^T$.
- For $\alpha \neq 0$, $(\alpha A)^\dagger = \alpha^{-1} A^\dagger$.

Example

- If $x \in \mathbb{R}$,

$$x^{\dagger} = \begin{cases} 0 & \text{if } x = 0; \\ x^{-1} & \text{otherwise.} \end{cases}$$

- If $x \in \mathbb{R}^n$,

$$x^{\dagger} = \begin{cases} 0^T & \text{if } x = 0; \\ x^T / x^T x & \text{otherwise.} \end{cases}$$

Summary

Norm

V is a vector space over $\mathbb{R}$ (or $\mathbb{C}$). $\|\cdot\| : V \times V \to \mathbb{R}$ is a norm if

1. $\|u\| \geq 0$; and $\|u\| = 0$ if and only if $u = 0$.
2. $\|\alpha u\| = |\alpha|\|u\|$.
3. $\|u + v\| \leq \|u\| + \|v\|$.

Inner product

V is a vector space over $\mathbb{R}$. $\langle\cdot,\cdot\rangle : V \to V \to \mathbb{R}$ is an inner product if

1. $\langle u,v\rangle = \langle v,u\rangle$
2. $\langle \alpha u+\beta v,w\rangle = \alpha\langle u,w\rangle + \beta\langle v,w\rangle$.
3. $\langle u,u\rangle \geq 0$; and $\langle u,u\rangle = 0$ if and only if $u=0$.

V is a vector space over $\mathbb{C}$. $\langle\cdot,\cdot\rangle : V \to V \to \mathbb{C}$ is a complex inner product if

1. $\langle u,v\rangle = \overline{\langle v,u\rangle}$
2. $\langle \alpha u+\beta v,w\rangle = \alpha\langle u,w\rangle + \beta\langle v,w\rangle$. $\langle w,\alpha u+\beta v\rangle = \overline{\alpha}\langle w,u\rangle + \overline{\beta}\langle w,v\rangle$
3. $\langle u,u\rangle \geq 0$; and $\langle u,u\rangle = 0$ if and only if $u=0$.

Cauchy-Schwarz inequlity and induced norm

V is an (complex) inner space. Then

$$|\langle u,v\rangle| \leq [\langle u,u\rangle]^{1/2}[\langle v,v\rangle]^{1/2}.$$

So an inner product induces a norm $\|v\| = \langle v,v\rangle^{1/2}$.

The adjoint of a linear operator

X,U two inner product spaces, $T : X \to U$ is linear. $\mathcal{X} = \{x_1, \cdots, x_n\}$, $\mathcal{U} = \{u_1, \cdots, u_m\}$ bases. Then there is a linear operator $L^* : U \to X$ which satisfies

$$\langle L(x), u \rangle_X = \langle x, L^*(u) \rangle_U .$$

Let $G_{ij} = \langle x_j, x_i \rangle_X$ and $M_{ij} = \langle u_j, L(x_i) \rangle_U$. Then $[L^*]_{\mathcal{U},\mathcal{X}} = G^{-1}B$.

Proof of the existence of the adjoint

Let $x = \sum \alpha_i x_i$ and $u = \sum \beta_j u_j$.

$$\langle T(x), u\rangle = \left\langle \sum_i \alpha_i T(x_i), \sum_j \beta_j u_j \right\rangle = \sum_i \sum_j \alpha_i \overline{\beta_j} \langle T(x_i), u_j\rangle$$
$$= \sum_i \alpha_i \overline{\sum_j \langle u_j, T(x_i)\rangle \beta_j} = \sum_i \alpha_i \overline{(M\beta)_i} = \alpha \cdot M\beta.$$

Since $[S(u)] = B\beta$, $S(u) = \sum_k (B\beta)_k x_k$.

$$\langle x, S(u)\rangle = \left\langle \sum_i \alpha_i x_i, \sum_k (B\beta)_k x_k \right\rangle = \sum_i \sum_k \alpha_i \overline{(B\beta)_k} \langle x_i, x_k\rangle$$
$$= \sum_i \alpha_i \overline{\sum_k \langle x_k, x_i\rangle (B\beta)_k} = \sum_i \alpha_i (GB\beta)_i = \alpha \cdot GB\beta.$$

The projection theorem

Let V be an inner product space and S be a finite dimensional subspace of V. For $v \in V$, there is a unique $\mathrm{proj}_S v = w \in S$ such that

$$\|v - w\| = \min_{z \in S} \{\|v - z\|\}.$$

Moreover, $\mathrm{proj}_S v = w$ if and only if $\langle v - w, z \rangle = 0$ for all $z \in S$. Let $\{u_1, \cdots, u_n\}$ be a basis for S, and $G_{ij} = \langle u_j, u_i \rangle$, $b_i = \langle v, u_i \rangle$. Then $\mathrm{proj}_S v = \sum x_i u_i$ where $x = G^{-1} b$.

Proof of the last formula

Let $w = \mathrm{proj}_S v = \sum x_i u_i$. Then $\langle v - w, z\rangle = 0$ for all $z \in S$ if and only if $\langle v - w, u_i\rangle$ for all i.

$$0 = \langle v - w, u_i\rangle = \left\langle v - \sum x_j u_j, u_i \right\rangle = \langle v, u_i\rangle - \sum_j x_j \langle u_j, u_i\rangle = b_i - (Gx)_i.$$

Thus x is a solution of $Gx = b$.

Orthogonal case

If $\{u_1, \cdots, u_n\}$ is an orthogonal set, the Gram matrix is the identity matrix I. Thus $x_i = b_i = \langle v, u_i\rangle$. Moreover $\langle \mathrm{proj}_S v, u_i\rangle = x_i \langle u_i, u_i\rangle$.

$$\mathrm{proj}_S v = \sum \frac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle} u_i.$$

The Gram-Schmidt process

For given linearly independent set $\{u_1, \cdots, u_n\}$, we can find an orthogonal set $\{\hat{u}_1, \cdots, \hat{u}_n\}$ such that

$$\hat{u}_1 = u_1, \hat{u}_{i+1} = u_{i+1} - \sum_{k=1}^{i} \frac{\langle u_{i+1}, u_k \rangle}{\langle u_k, u_k \rangle} u_k.$$

If we denote $S_k = \operatorname{span}\{\hat{u}_1, \cdots, \hat{u}_k\}$,

$$\hat{u}_{i+1} = u_{i+1} - \operatorname{proj}_{S_k} u_{i+1}.$$

The least square solution

Let $A \in \mathbb{C}^{m\times n}$ and $y \in \mathbb{C}^m$. We can define a least square solution to $Ax = y$ by

$$\|Ax - y\|_2^2 = \min_{z\in\mathbb{C}^m} \{\|Az - y\|_2^2\}.$$

x is a LSS to $Ax = y$ iff $\mathrm{proj}_{\mathrm{col}(A)} y = Ax$
iff $\langle Ax - y, Az\rangle = 0$ for all $z \in \mathbb{C}^n$
iff $\langle A^*Ax - A^*y, z\rangle = 0$ for all $z \in \mathbb{C}^n$
iff $A^*Ax = A^*y$.

The minimum norm least square solution

Let $\hat{x}$ be a LSS to $Ax = y$ and $\overline{x} = \mathrm{proj}_{\mathrm{col}(A^*)}\hat{x}$. Then for all $z \in \mathbb{C}^m$,

$$0 = \langle \hat{x} - \overline{x}, A^*z\rangle = \langle A\hat{x} - A\overline{x}, z\rangle$$

So $A\hat{x} = A\overline{x}$ and this implies $\overline{x}$ is also a LSS to $Ax = y$. Furthermore, $\overline{x}$ is unique.

Orthogonal projection

Let $P(v) = \mathrm{proj}_S v$. Then P is linear, $P^2 = P$, and $P^* = P$.

Proof.

For $s \in S$ and $\alpha \in \mathbb{C}$,

$$\langle (v+w) - P(v) - P(w), s\rangle = \langle v - P(v), s\rangle + \langle w - P(w), s\rangle = 0.$$
$$\langle \alpha v - \alpha P(V), s\rangle = \alpha\langle v - P(v), s\rangle = 0.$$

So $P(v+w) = P(v) + P(w)$ and $P(\alpha v) = \alpha P(v)$.
For $v, s \in S$, $\langle v - P(v), s\rangle = 0$ implies $P(v) = v$. Thus $P^2(v) = P(v)$ for all $v \in V$.
Finally, for all $u, v \in V$,

$$\begin{aligned}\langle P(v), w\rangle - \langle v, P(w)\rangle &= \langle P(v), w\rangle - \langle P(v), P(w)\rangle + \langle P(v), P(w)\rangle - \langle v, P(w)\rangle \\ &= \langle P(v), w - P(w)\rangle - \langle v - P(v), w\rangle = 0\end{aligned}$$

$\square$

Orthogonal projection

Suppose $P : V \to V$ is linear map such that $P^2 = P$ and $P^* = P$. Then P is an orthogonal projection to some subspace of V.

Proof.

Let $S = \mathcal{R}(P)$. For $v \in V$ and $P(w) \in S$,

$$\begin{aligned} \langle v - P(v), P(w) \rangle &= \langle P^*(v - P(v)), w \rangle \\ &= \langle P(v) - P^2(v), w \rangle \\ &= \langle P(v) - P(v), w \rangle = 0 \end{aligned}$$

So $P(v) = \mathrm{proj}_S v$. □

Pseudoinverse

Thus we can define a function $S : \mathbb{C}^m \to \mathrm{col}(A^*)$ by $S(y) = \overline{x}$. It is a routine proof that S is linear. Let $A^\dagger$ be the corresponding matrix of S. $A^\dagger$ is the pseudoinverse of A.

$AA^\dagger$ is an orthogonal projection to $\mathrm{col}(A)$ and $A^\dagger A$ is an orthogonal projection to $\mathrm{col}(A^*)$.

We can show that
$B = A^\dagger$ if and only if $ABA = A$, $BAB = B$, $BA = (BA)^*$ and $AB = (AB)^*$.

Ch7. The spectral theory of symmetric matrices

Recall

- $A \in \mathbb{R}^{n\times n}$ is called *symmetric* if $A^T = A$.
- A is diagonalizable if there is invertible matrix X and diagonal matrix D such that $A = XDX^{-1}$.

For convenience, if we say a matrix A is symmetric, we may assume $A \in \mathbb{R}^{m\times n}$.
The goal of this chapter :

- every symmetric (or Hermitian) matrix is diagonalizable
- the eigenvalues of symmetric (or Hermitian) matrix are real
- we can choose "nice" eigenvectors of a symmetric (or Hermitian)

diagonalizable

Let $A \in \mathbb{C}^{n\times n}$. Then A has eigenvalues $\lambda_1, \cdots, \lambda_k$ such that

$$\mathrm{m.alg}(\lambda_1) + \cdots + \mathrm{m.alg}(\lambda_k) = n.$$

For each i, $\mathrm{m.geo}(\lambda_i) \le \mathrm{m.alg}(\lambda_i)$.

A is diagonalizable if and only if $\mathrm{m.geo}(\lambda_i) = \mathrm{m.alg}(\lambda_i)$ for all i.

TFAE

- $E_\lambda(A) = G_\lambda(A)$.
- $\mathrm{m.geo}(\lambda) = \mathrm{m.alg}(\lambda)$.
- $\mathcal{N}(A - \lambda I) = \mathcal{N}(A - \lambda I)^2$.
- $\mathcal{N}(A - \lambda I) \cap \mathrm{col}(A - \lambda I) = \{0\}$.

7.1 The spectral theorem for symmetric matrices

Theorem (328)

Let $A \in \mathbb{R}^{n\times n}$ be symmetric, and let $\lambda \in \mathbb{C}$ be an eigenvalue of A. Then $\lambda \in \mathbb{R}$ and there exists an eigenvector $x \in \mathbb{R}^n$ corresponding to λ.

Theorem (329)

Let $A \in \mathbb{R}^{n\times n}$ be symmetric, and let $\lambda_1, \lambda_2 \in \mathbb{R}$ be distinct eigenvalues of A, and let $x_1, x_2 \in \mathbb{R}^n$ be eigenvectors corresponding to λ_1, λ_2, respectively. Then x_1 and x_2 are orthogonal.

Definition

Let $Q \in \mathbb{R}^{n \times n}$. We say that Q is *orthogonal* if and only if $Q^T = Q^{-1}$.

Equivalently, $Q_i \cdot Q_j = \delta_{ij}$ where δ_{ij} is Kronecker delta. Thus $\{Q_1, \cdots, Q_n\}$ is <u>orthonormal</u>.

Theorem (331)

Let $A \in \mathbb{R}^{n \times n}$ be symmetric and let $\lambda \in \mathbb{R}$ be an eigenvalue of A. Then the geometric multiplicity of λ equals the algebraic multiplicity of λ.

Proof

Induction on n. Assume the result hold for all matrices of dimension $(n-1) \times (n-1)$. Suppose $\lambda \in \mathbb{R}$ is an e.val of A of $\mathrm{m.alg}(\lambda) = k \geq 1$. Let $x \in \mathbb{R}^n$ be an e.vec of A corr to λ. We may assume $\|x\| = 1$ and can find an orthonormal basis $\{x_1 = x, x_2, \cdots, x_n\}$ for $\mathbb{R}^n$. Define $X = [x_1 | \cdots | x_n]$. Then X is orthogonal.

Proof

Define $B = X^T AX$.

$$B = \begin{bmatrix} x_1 \cdot Ax_1 & \cdots & x_1 \cdot Ax_n \\ \vdots & \ddots & \vdots \\ x_n \cdot Ax_1 & \cdots & x_n \cdot Ax_n \end{bmatrix}$$

Since $Ax_1 = \lambda x_1$, $x_i \cdot Ax_1 = 0$. Also $x_1 \cdot Ax_i = Ax_1 \cdot x_i = 0$. Thus

$$B = \begin{bmatrix} x_1 \cdot Ax_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n \cdot Ax_n \end{bmatrix}$$

Write $C = B^{(1,1)}$;

$$B = \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & C \end{array}\right]$$

Since B is symmetric, so is C.

Proof

Then λ is an eigenvalue of C of $\mathrm{m.alg} = k-1$. By I.H, C has $k-1$ e.vec $u_2, \cdots, u_k$(may assume it is orthonormal). Define

$$z_i = \left[\begin{array}{c} 0 \\ \hline u_i \end{array}\right]$$

Then

$$Bz_i = \left[\begin{array}{c|c} \lambda & 0 \\ \hline 0 & C \end{array}\right] \left[\begin{array}{c} 0 \\ \hline u_i \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline \lambda u_i \end{array}\right] = \lambda z_i$$

Finally, $AXz_i = \lambda z_i$ implies $Xz_2, \cdots, Xz_n$ are all e.vec of A corr to λ.

Corollary (332)

Let $A \in \mathbb{R}^{n \times n}$ be symmetric. Then there exists an orthogonal matrix $X \in \mathbb{R}^{n \times n}$ and a diagonal matrix $D \in \mathbb{R}^{n \times n}$ such that $A = XDX^T$.

Definition

Let $A \in \mathbb{R}^{n \times n}$ be symmetric. We say that A is symmetric positive definite (SPD) if and only if

$$x \cdot (Ax) > 0 \text{ for all } x \in \mathbb{R}^n - \{0\}.$$

A is positive semidefinite if and only if

$$x \cdot (Ax) \geq 0 \text{ for all } x \in \mathbb{R}^n.$$

In other words, if A is SPD, then "$x \cdot Ax = 0$ iff $x = 0$." If A is just positive semidefinite, there may be nonzero x such that $x \cdot Ax = 0$.

Theorem (334)

Let $A \in \mathbb{R}^{n\times n}$ be symmetric.

- *A is SPD if and only if all of the e.val of A are positive.*
- *A is positive semidefinite if and only if all of the e.val of A are nonnegative.*

Theorem (335)

Let X be an inner product space over $\mathbb{R}$ and let $\{x_1, \cdots, x_n\}$ be a basis for X. Then the Gram matrix G is SPD.

Corollary (336)

Let $A \in \mathbb{R}^{m\times n}$ be nonsingular. Then $A^T A$ is SPD.

Theorem (337)

Let $A \in \mathbb{C}^{n \times n}$ *be Hermitian, and let* λ *be e.val of* A. *Then* $\lambda \in \mathbb{R}$.

Theorem (338)

Let $A \in \mathbb{C}^{n \times n}$ *be Hermitian, and let* λ *be e.val of* A. *Then* $\mathrm{m.geo}(\lambda) = \mathrm{m.alg}(\lambda)$.

Theorem (339)

Let $A \in \mathbb{C}^{n \times n}$ *be Hermitian, and let* λ_1, λ_2 *be e.val of* A, *and* $x_1, x_2 \in \mathbb{C}^n$ *be e.vec corr to* λ_1, λ_2*,respectively. Then* x_1 *and* x_2 *are orthogonal.*

Definition

$U \in \mathbb{C}^{n \times n}$ is called <u>unitary</u> if and only if $U^* = U^{-1}$.

Theorem (340)

Let $A \in \mathbb{C}^{n \times n}$ *be Hermitian. Theb there exists a unitary matrix* $X \in \mathbb{C}^{n \times n}$ *and a diagonal matrix* $D \in \mathbb{R}^{n \times n}$ *such that* $A = XDX^*$.

Definition

Let $A \in \mathbb{C}^{n\times n}$ be Hermitian. We say that A is positive definite if and only if

$$x \cdot (Ax) > 0 \text{ for all } x \in \mathbb{C}^n - \{0\}.$$

A is positive semidefinite if and only if

$$x \cdot (Ax) \geq 0 \text{ for all } x \in \mathbb{C}^n.$$

Theorem (341)

Let $A \in \mathbb{C}^{n\times n}$ be Hermitian.

- *A is positive definite if and only if all of the e.val of A are positive.*
- *A is positive semidefinite if and only if all of the e.val of A are nonnegative.*

Theorem (343)

Let X be an inner product space over $\mathbb{C}$ and let $\{x_1, \cdots, x_n\}$ be a basis for X. Then the Gram matrix G is Hermitian positive definite.

Corollary (344)

Let $A \in \mathbb{C}^{m \times n}$ be nonsingular. Then $A^* A$ is Hermitian and positive definite.

Ex7.1.1

Let $A \in \mathbb{R}^{m \times n}$. Prove that $A^T A$ is positive semidefinite.

Ex7.1.2

Prove that a symmetric matrix $A \in \mathbb{R}^{n \times n}$ is positive semidefinite if and only if all the eigenvalues of A are nonnegative.

Ex7.1.3

Let $A \in \mathbb{R}^{n\times n}$ be symmetric and positive definite. Define $\langle\cdot,\cdot\rangle$ by

$$\langle x, y\rangle = x \cdot Ay.$$

Prove that $\langle\cdot,\cdot\rangle$ is an inner product on $\mathbb{R}^n$.

Ex7.1.4

Let $U \in \mathbb{R}^{n\times n}$ be an orthogonal matrix. Prove that multiplication by U preserves norms and dot products:

$$\|Ux\|_2 = \|x\|_2$$
$$(Ux)\cdot(Uy) = x\cdot y$$

Ex7.1.5

Suppose multiplication by $A \in \mathbb{R}^{n \times n}$ preserves dot products:

$$(Ax) \cdot (Ay) = x \cdot y.$$

Does A have to be orthogonal? Prove or disprove.

Ex7.1.6

Suppose multiplication by $A \in \mathbb{R}^{n\times n}$ preserves dot norms:

$$\|Ax\|_2 = \|x\|_2.$$

Does A have to be orthogonal? Prove or disprove.

Ex7.1.7

Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian positive definite matrix. Prove that there exists a Hermitian positive definite matrix $B \in \mathbb{C}^{n\times n}$ such that $B^2 = A$. (The matrix B is calles the square root of A and is denoted $A^{1/2}$.)

Ex7.1.8

Let X be a finite-dimensional inner product space over $\mathbb{C}$, and let $T : X \to X$ be a linear operator. We say T is self-adjoint if $T^* = T$. Suppose T is self-adjoint. Prove:

(a) Every e.val of T is real.

(b) e.vec of T corr to distinct e.val are orthogonal.

Ex7.1.9

X finite-dimensional inner product space over $\mathbb{R}$ with basis $\mathcal{X} = \{x_1, \cdots, x_n\}$, and $T : X \to X$ is a self-adjoint linear operator. Let $A = [T]_{\mathcal{X},\mathcal{X}}$ and G be the Gram matrix for $\mathcal{X}$, and $B \in \mathbb{R}^{n\times n}$ by

$$B = G^{1/2} A G^{-1/2}.$$

(a) Prove that B is symmetric.

(b) Since A and B are similar, they have the same eigenvalues and there is a simple relationship between their eigenvectors. What is this relationship?

(c) Use the fact that there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of B to prove that there is an orthonormal basis of X consisting of eigenvectors of T.

Ex7.1.9(a)

Prove that B is symmetric.

Ex7.1.9(b)

Since A and B are similar, they have the same eigenvalues and there is a simple relationship between their eigenvectors. What is this relationship?

Ex7.1.9(c)

Use the fact that there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of B to prove that there is an orthonormal basis of X consisting of eigenvectors of T.

Some properties of dual spaces

Matrix and dual space

Let $V = \mathbb{R}^n$. Every vector $v \in V$ is a matrix of size $n \times 1$. Since every matrix defines a linear operator, we can correspond v to a linear operator.

$$v : \mathbb{R} \to \mathbb{R}^n.$$

Similarly, $v^T \in \mathbb{R}^{1 \times n}$.

$$v^T : \mathbb{R}^n \to \mathbb{R}.$$

So v^T is a dual vector of v and $\mathbb{R}^{1 \times n} \cong \mathcal{L}(\mathbb{R}^n, \mathbb{R}) = V^*$.

Dot product and dual vector space

Let $u, v \in V = \mathbb{R}^n$. We know that $u \cdot v = v^T u$. So the map $v \mapsto v^T = (\text{-}) \cdot v$ is a linear map from V to V^*. Moreover e_i^T is a basis for V^* because $e_i^T e_j = \delta_{ij}$. Similarly if $\{v_1, \cdots, v_n\}$ is a basis for V, then $\{v_1^T, \cdots, v_n^T\}$ is a basis for V^*.

Inner product and dual vector space

Let V be a finite dimensional inner product space over $\mathbb{R}$. The correspondence $L : v \mapsto \langle \cdot, v \rangle$ is a linear map from V to V^*. On the other hand $\mathcal{B} = \{v_1, \cdots, v_n\}$ is a basis for V, then $\mathcal{B}^* = \{v_1^*, \cdots, v_n^*\}$ is a basis for V^* where $v_i^*(v_j) = \delta_{ij}$. Now compute $M = [L]_{\mathcal{B},\mathcal{B}^*}$.

$$M = ([L(v_1)]| \cdots |[L(v_n)])$$

$$L(v_i) = \sum \alpha_{ij} v_j^* = \langle \cdot, v_i \rangle.$$

Since $L(v_i)(v_j) = \langle v_j, v_i \rangle = \alpha_{ij} v_j^*(v_j) = \alpha_{ij}$, $\alpha_{ij} = (G_{\mathcal{B}})_{ij}$. Hence $[L] = G_{\mathcal{B}}$.

Continued

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{G_{\mathcal{B}}} & \mathbb{R}^n \\ E_{\mathcal{B}} \uparrow & & \uparrow E_{\mathcal{B}^*} \\ V & \xrightarrow{L} & V^* \end{array}$$

Theorem (Dual linear map)

Let U and V be two finite dimensional vector spaces over F. Suppose $L : U \to V$ is a linear map. Then there is a linear map $S : V^* \to U^*$ such that $S(f) = f \circ L$. Write $S = \hat{L}$.

Proof.

$$
\begin{tikzcd}
V \arrow[r, "f"] & \mathbb{R} \\
U \arrow[u, "L"] \arrow[ur, "f \circ L"'] &
\end{tikzcd}
.
$$

Let $f \in V^*$. Then $f \circ L : U \to \mathbb{R}$ is a linear map. So $S(f)$ is well-defined map. Moreover $S(f + g) = (f + g) \circ L = f \circ L + g \circ L = S(f) + S(g)$ and $S(\alpha f) = (\alpha f) \circ L = \alpha(f \circ L) = \alpha S(f)$. Hence S is a linear map. $\square$

Theorem

Let U, V, and W be finite dimensional vector spaces over F. Suppose $T : U \to V$ and $S : V \to W$. Then $\widehat{(ST)} = \hat{T}\hat{S}$.

Proof.

$$\widehat{(ST)}(f) = fST = (\hat{S})T = \hat{T}(\hat{S}(f)).$$

$$U \xrightarrow{T} V \xrightarrow{S} W$$

$$U^* \xleftarrow[\hat{T}]{} V^* \xleftarrow[\hat{S}]{} W^*$$

$\square$

Theorem

Let U and V be finite dimensional inner prodcut spaces over $\mathbb{R}$. Then we have isomorphisms, $\phi^U : u \mapsto \langle \cdot, u \rangle_U$ and $\phi^V : v \mapsto \langle \cdot, v \rangle_V$. Let $L : U \to V$ be a linear map. $L^* : V \to U$ is a linear map such that $\langle L(u), v \rangle_V = \langle u, L^*(v) \rangle_U$. Then $\phi^U L^* = \hat{L} \phi^V$.

$$\begin{tikzcd} U \arrow[d, "\phi^U"'] & V \arrow[l, "L^*"'] \arrow[d, "\phi^V"] \\ U^* & V^* \arrow[l, "\hat{L}"] \end{tikzcd}$$

Proof.

$$(\phi^U L^*(v))(u) = \langle v, L^*(u) \rangle_U = \langle L(u), v \rangle_V$$
$$(\hat{L} \phi^V(v))(u) = (\phi^V(v))(L(U)) = \langle L(u), v \rangle_V.$$

$\square$

Matrix for dual maps

$$
\begin{tikzcd}
\mathbb{R}^m \arrow[ddd, "G_{\mathcal{U}}"'] & & & \mathbb{R}^n \arrow[lll, "B"'] \arrow[ddd, "G_{\mathcal{V}}"] \\
& U \arrow[ul, "E_{\mathcal{U}}"] \arrow[d, "\phi^U"'] & V \arrow[l, "L^*"'] \arrow[ur, "E_{\mathcal{V}}"'] \arrow[d, "\phi^V"] & \\
& U^* \arrow[dl, "E_{\mathcal{U}^*}"'] & V^* \arrow[l, "\hat{L}"] \arrow[dr, "E_{\mathcal{V}^*}"] & \\
\mathbb{R}^m & & & \mathbb{R}^n \arrow[lll, "A"]
\end{tikzcd}
$$

Compute $A = [\hat{L}]_{\mathcal{V}^*,\mathcal{U}^*}$. Let $M_{ij} = \langle v_j, L(u_i)\rangle$. We already show that $B = (G_{\mathcal{U}})^{-1}M$.

$$
\begin{aligned}
[\hat{L}]_{\mathcal{V}^*,\mathcal{U}^*} &= A = G_{\mathcal{U}}B(G_{\mathcal{V}})^{-1} = M(G_{\mathcal{V}})^{-1} \\
&= ((G_{\mathcal{V}}^{-1})^T M^T)^T = (G_{\mathcal{V}}^{-1}M^T)^T = [L]_{\mathcal{U},\mathcal{V}}^T.
\end{aligned}
$$

The last equality is followed by $(A^T)^{-1} = (A^{-1})^T$, $G^T = G$, and $(M^T)_{ij} = \langle u_j, L^*(v_i)\rangle$

Theorem

Let V be a vector sapce over F. Let $v \in V$. If $f(v) = 0$ for all $f \in V^*$, then $v = 0$.

Corollary

If $f(v_1) = f(v_2)$ for all $f \in V^*$, then $v_1 = v_2$.

Proof.

If $v \neq 0$, we can find a basis for V containing v. Then there is a linear map $f : V \to \mathbb{R}$ such that $f(v) = 1$. Hence if $f(v) = 0$ for all $f \in V^*$, $v = 0$. □

Remark

If V is an inner product space over $\mathbb{R}$, $\langle v, w \rangle = 0$ for all $w \in V$ implies $v = 0$. Let V be a vector space over F and let $\langle \cdot, \cdot \rangle : V \times V^* \to F$ be an evalutation map. Above theorem says $\langle v, f \rangle = 0$ for all $f \in V^*$, $v = 0$.

Exercises in Dummit 11.3

Let S be any subset of V^* for some finite dimensional vector space V. Define $\mathrm{Ann}(S) = \{v \in V \mid f(v) = 0 \text{ for all } f \in S\}$. We call $\mathrm{Ann}(S)$ the annihilator of S in V.

(a) $\mathrm{Ann}(S)$ is a subspace of V.

(b) Let W_1, W_2 be subspaces of V^*. Then $\mathrm{Ann}(W_1 + W_2) = \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2)$ and $\mathrm{Ann}(W_1 \cap W_2) = \mathrm{Ann}(W_1) + \mathrm{Ann}(W_2)$.

(c) $W_1 = W_2$ if and only if $\mathrm{Ann}(W_1) = \mathrm{Ann}(W_2)$.

(d) $\mathrm{Ann}(S) = \mathrm{Ann}(\mathrm{span}(S))$.

(e) Let $\{v_1, \cdots v_n\}$ be a basis for V. If $S = \{v_1^*, \cdots, v_k^*\}$ for some $k \leq n$, then $\mathrm{Ann}(S) = \mathrm{span}\{v_{k+1}, \cdots, v_n\}$.

(f) If W^* is a subspace of V^*, then $\dim \mathrm{Ann}(W^*) = \dim V - \dim W^*$.

Remark

If $A \subset B \subset V^*$,

$$\mathrm{Ann}(A) \supset \mathrm{Ann}(B).$$

Recall

If V is finite dimensional, $V \to V^{**}$ by $v \mapsto \langle v, \cdot \rangle$ is an isomorphism.

(a)

$\mathrm{Ann}(S)$ is a subspace of V.

Proof.

Let $v, w \in \mathrm{Ann}(S)$. For $f \in S$, $f(0) = 0$, $f(v+w) = f(v) + f(w) = 0 + 0 = 0$, $f(\alpha v) = \alpha f(v) = 0$. Hecne $0, v + w, \alpha v \in \mathrm{Ann}(S)$. $\square$

(b)

Let W_1, W_2 be subspaces of V^*. Then $\mathrm{Ann}(W_1 + W_2) = \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2)$ and $\mathrm{Ann}(W_1 \cap W_2) = \mathrm{Ann}(W_1) + \mathrm{Ann}(W_2)$.

Proof.

Since $W_1 + W_2 \supset W_i$, $\mathrm{Ann}(W_1 + W_2) \subset \mathrm{Ann}(W_i)$. So $\mathrm{Ann}(W_1 + W_2) \subset \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2)$.

Let $f_1 + f_2 \in W_1 + W_2$. Let $v \in \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2)$. Then $(f_1 + f_2)(v) = f_1(v) + f_2(v) = 0$. So $\mathrm{Ann}(W_1 + W_2) \supset \mathrm{Ann}(W_1) \cap \mathrm{Ann}(W_2)$

Since $W_1 \cap W_2 \subset W_i$, $\mathrm{Ann}(W_1 \cap W_2) \supset \mathrm{Ann}(W_1) + \mathrm{Ann}(W_2)$. (continued)

Proof.

To show $\mathrm{Ann}(W_1 \cap W_2) \subset \mathrm{Ann}(W_1) + \mathrm{Ann}(W_2)$, choose a basis for $W_1 \cap W_2$ and extend it to a basis for $W_1 + W_2$ and one more to V^*, say $\mathcal{B}$. Let $v \in \mathrm{Ann}(W_1 \cap W_2)$. Define a map $E_i : V^* \to F$ by

$$E_1(f) = \begin{cases} 0 & \text{for } f \in \mathcal{B} \cap W_1 \\ f(v) & \text{otherwise} \end{cases}$$

$$E_2(f) = \begin{cases} f(v) & \text{for } f \in \mathcal{B} \cap W_1 \\ 0 & \text{otherwise.} \end{cases}$$

Since V is finite dimensional, $V \cong V^{**}$. Thus we can find $v_i \in V$ such that $v_i \mapsto E_i$. Then for $f_i \in \mathcal{B} \cap W_i$, $f_1(v_j) = \langle v_j, f_i \rangle = E_j(f_i)$. Since $E_1(f_1) = 0 = f_1(v_1)$ and $E_2(f_2) = 0 = f_2(v_2)$, $v_i \in W_i$. Now for $f \in \mathcal{B}$, $f(v_1 + v_2) = f(v_1) + f(v_2) = E_1(f) + E_2(f) = f(v)$. Thus for all $f \in V^*$, $f(v_1 + v_2) = f(v)$. Hence $v = v_1 + v_2$, as desired. □

(c)

$W_1 = W_2$ if and only if $\mathrm{Ann}(W_1) = \mathrm{Ann}(W_2)$.

Proof.

Claim) If $W_1 \subset W_2 \subset V^*$ and $\mathrm{Ann}(W_1) = \mathrm{Ann}(W_2)$, then $W_1 = W_2$.

Suppose not. Then there is $g \in W_2 - W_1$. We can find a basis $\mathcal{B}$ containing g such that $\mathcal{B} \cap W_1$ is a basis for W_1. We defind a map $E : V^* \to F$ by

$$E(f) = \begin{cases} 1 & \text{if } f = g \\ 0 & \text{if } f \neq g \end{cases}$$

Now let $w \in V$ whose image is E, i.e. $\langle w, f \rangle = E(f)$. Since $\langle w, f \rangle = E(f) = 0$ for $f \in W_1$, $w \in \mathrm{Ann}(W_1)$. But $\langle w, g \rangle = E(g) = 1$. So $\mathrm{Ann}(W_1) \neq \mathrm{Ann}(W_2)$ (contradiction.)

Using the claim, $\mathrm{Ann}(W_1 \cap W_2) = \mathrm{Ann}(W_1) + \mathrm{Ann}(W_2) = \mathrm{Ann}(W_i)$ implies $W_1 \cap W_2 = W_i$. Hence $W_1 = W_2$. $\square$

(d)

$\text{Ann}(S) = \text{Ann}(\text{span}(S))$.

Proof.

Clearly $\text{Ann}(S) \supset \text{Ann}(\text{span}(S))$. Let $\sum \alpha_i f_i \in \text{span}(S)$ and $v \in \text{Ann}(S)$.

$$(\sum \alpha_i f_i)(v) = \sum \alpha_i f_i(v) = 0.$$

Hence $v \in \text{Ann}(\text{span}(S))$. $\square$

(e)

Let $\{v_1, \cdots v_n\}$ be a basis for V. If $S = \{v_1^*, \cdots, v_k^*\}$ for some $k \leq n$, then $\mathrm{Ann}(S) = \mathrm{span}\{v_{k+1}, \cdots, v_n\}$.

Proof.

Since $v_i^*(v_j) = \delta_{ij}$, $v_1, \cdots, v_k \notin \mathrm{Ann}(S)$ and $v_{k+1}, \cdots, v_n \in \mathrm{Ann}(S)$. $\square$

(f)

If W^* is a subspace of V^*, then $\dim \mathrm{Ann}(W^*) = \dim V - \dim W^*$.

Proof.

Claim) There is a linearly independent set $\{v_1, \cdots, v_k\} \subset V$ such that $\{v_1^*, \cdots, v_k^*\} \subset V^*$ is a basis for W^*.

Let $\{f_1, \cdots, f_k\}$ be a basis for W^* and extend to $\{f_1, \cdots, f_n\}$. Define $E_i : V^* \to F$ by $E_i(f_j) = \delta_{ij}$. Then $\{E_1, \cdots, E_k\}$ is linearly independent. Let $v_i \in V$ such that $\langle V_i, \cdot \rangle = E_i$. Then $f_i(v_j) = \langle v_j, f_i \rangle = E_j(f_i) = \delta_{ij} = \langle v_j, v_i^* \rangle$. Thus $f_i = v_i^*$. So $\{v_1^*, \cdots, v_k^*\}$ is a basis for W^*.

Now using (e), $\dim \mathrm{Ann}(W) = \dim V - \dim W^*$. $\square$

Annihilator vs orthogonal complement

Let V be an finite dimensional inner product space. For $S \subset V$, $S^{\perp} = \{v \in V \mid \langle v, s \rangle = 0 \text{ for all } s \in S\}$.

Let $v \in V$. If $v \neq 0$, we can extend $\{v\}$ to an orthogonal basis $\mathcal{B}$ for V. Then $v^*(w) = \delta_{vw} = \langle w, v \rangle / \langle v, v \rangle$ for $w \in \mathcal{B}$. Let $S^* = \{s^* \mid s \in S\}$. Note that $0^* = 0$.
For $v \in S^{\perp}$ and for $s \neq 0$, $s^*(v) = \langle v, s \rangle / \langle s, s \rangle = 0$. So $S^{\perp} \subset \mathrm{Ann}(S^*)$.
Conversely, for $v \in \mathrm{Ann}(S^*)$, $\langle v, s \rangle = \langle s, s \rangle s^*(v) = 0$. So $S^{\perp} \supset \mathrm{Ann}(S^*)$.

Summary

Suppose V is a finite dimensional inner product space. Then $V \cong V^*$. Let $\langle \cdot, \cdot \rangle_I$ be an inner product and $\langle \cdot, \cdot \rangle_E$ be an evaluation map. The map $v \mapsto \langle \cdot, v \rangle_I$ is an isomorphism from V and V^*.
Let $\{v_1, \cdots, v_n\}$ be an orthonormal basis for V. Then

$$\langle v_j, v_i^* \rangle_E = v_i^*(v_j) = \langle v_j, v_i \rangle_I$$

Thus we can identity v with v^* and $\langle \cdot, \cdot \rangle_I$ with $\langle \cdot, \cdot \rangle_E$.

Or an evaluation map is a generalization of an inner product.

The End