1 Review: Minimisation

Recall that $V = W \oplus W^{\perp}$.

Therefore, if $x \in V$, then $x = x_W + x_{W^{\perp}}$, where x_W is the vector in W closest to $x \in V$, the unique vector in W such that $x - x_W \in W^{\perp}$.

2 Adjoints

Theorem 2.1

Let V, W be finite dimensional inner product spaces. Let $T \in \text{Hom}(V, W)$.

Then there exists a unique homomorphism $T^* \in \text{Hom}(W, V)$ such that $\langle T(x), y\rangle_W = \langle x, T^*y\rangle_V$ for all $x \in V$ and $y \in W$.

Proof. Fix $y \in W$. Note that the function $f_y : W \to \mathbb{F}$, where $x \mapsto \langle x, y\rangle_W$ is linear.

Lemma 2.2

If $f \in \text{Hom}(W, V)$, then $f = f_y$ for a unique $y \in W$, i.e. $f(x) = \langle x, y\rangle_W$.

Fix $y \in W$.

Consider $f : V \to \mathbb{F}$, with $x \mapsto \langle Tx, y\rangle_W$.

This is linear: f is in the composition $f_y \circ T$ of two linear functions.

By Lemma 2.2, there exists a unique vector $T^*y \in V$ such that $f(x) = \langle x, T^*y\rangle_V$ for all $x \in V$. Thus, we define a function $T^* : W \to V$ such that Theorem 2.1 holds, and thus $f(x) = \langle Tx, y\rangle$.

Claim. T^* is unique.

Fix $y \in W$. Then $T^*y \in V$ is unique by the uniqueness of a vector given by Lemma 2.2.

Claim. T^* is linear.

Observe the following:

$$\begin{aligned}
\langle x, T^*(cy_1 + y_2)\rangle_V &= \langle Tx, cy_1 + y_2\rangle_W && (1)\\
&= \bar{c}\langle Tx, y_1\rangle_W + \langle Tx, y_2\rangle_W && (2)\\
&= \bar{c}\langle x, T^*y\rangle_V + \langle x, T^*y\rangle_V && (3)\\
&= \langle x, cT^*y_1 + T^*y_2\rangle && (4)
\end{aligned}$$

Therefore, $T^*(cy_1 + y_2) = cT^*y_1 + T^*y_2$

□

Remark 2.3. $\langle y, Tx\rangle_W = \langle T^*y, x\rangle_V$

Remark 2.4. If V is not finite-dimensional, then T^* may not exist.

Theorem 2.5

Let β be an orthonormal basis of V, a finite-dimensional vector space.

Let γ be an orthonormal basis of W, a finite-dimensional vector space.

Then $[T^*]_\gamma^\beta = ([T]_\beta^\gamma)^*$.

Proof. Let $A = [T]_\beta^\gamma$ and $B = [T^*]_\gamma^\beta$, where $\beta = (v_1, \ldots, v_n)$ and $\gamma = (w_1, \ldots, w_m)$. Therefore,

$$B_{ij} = \langle T^* w_j, v_i \rangle_V = \langle w_j, T v_i \rangle = \overline{\langle T v_i, w_j \rangle} = \overline{A_{ji}}.$$

Therefore, $B = A^*$. □

Theorem 2.6 a) If $T \in \mathrm{Hom}(V, W)$, $U \in \mathrm{Hom}(V, W)$, and $c \in \mathbb{F}$, then $(cT+U)^* = \overline{c}T^* + U^*$.

b) If $T : V \to W$ and $U : W \to Z$, where V, W, Z are finite-dimensional vector spaces, then $(UT)^* = T^* U^*$.

c) $T^{**} = T$

d) $I^* = I$, if the inner product is the same.

Proof. 2.6, b) and c) are left as exercises.

a) can be deduced as follows, for all $x \in V$ and $y \in Z$:

$$\langle x, (TU)^* y \rangle = \langle TUx, y \rangle_Z = \langle Ux, T^* y \rangle_W = \langle x, U^* T^* y \rangle_V$$

□

Corollary 2.7

If $A \in M_{m \times n}(\mathbb{F})$, then $L_A : \mathbb{F}^n \to \mathbb{F}^m$ and $(L_A)^* = L_{A^*}$, where standard inner products are used.

Moreover,

a) $(cA + B)^* = \overline{c}A^* + B^*$

b) $(AB)^* = B^* A^*$

c) $A^{**} = A$

Proof. Let $L_A : \mathbb{F}^n \to \mathbb{F}^m$, and let β, γ be standard bases for $\mathbb{F}^n, \mathbb{F}^m$ respectively.

Therefore, $[L_A^*]_\gamma^\beta = ([L_A]_\beta^\gamma)^* = A^* = [L_{A^*}]_\gamma^\beta$. Therefore, $L_A^* = L_{A^*}$.

Then we can deduce (a)-(d) from Theorem 2.6.

□

3 Least Square Approximation

Let $A \in M_{m\times n}(\mathbb{F})$, $y \in \mathbb{F}^m$.

Problem. Devise a method to find $x_0 \in \mathbb{F}^n$ such that $\|Ax_0 - y\| \leq \|Ax - y\|$ for all $x \in \mathbb{F}^n$.

Lemma 3.1

$\langle Ax, y\rangle_{F^n} = \langle x, A^*y\rangle_{F^m}$.

Proof. Apply Corollary 2.7. □

Lemma 3.2

rank $A^*A =$ rank A

Proof. It is enough to show that ker $A^*A =$ ker A.

If $Ax = 0$, then $A^*Ax = 0$, and hence ker $A \subseteq$ ker A^*A.

If $A^*Ax = 0$, then $\langle x, A^*Ax\rangle = \langle Ax, Ax\rangle = \|Ax\|^2$. Therefore, $Ax = 0$, and hence ker $A^*A \subseteq A$, and thus ker $A^*A =$ ker A. □

Theorem 3.3

a) $\exists x_0 \in T^n$ such that $\|Ax_0 - y\| \leq \|Ax - y\|$ for all $x \in \mathbb{F}^n$.

b) x_0 satisfies (a) $\Leftrightarrow$ $A^*Ax_0 = A^*y$.

c) If rank $A = n$, then A^*A is invertible, so there's a unique x_0 in (a).

Proof.

a) Let $W = \text{Im}(A)$. We know that there exists w_0 in any W such that $\|w_0 - y\| \leq \|w - y\|$ for all $W \in W$.

   Write $w_0 = Ax_0$ for some x_0. Therefore, $\|Ax_0 - y\| \leq \|Ax - y\|$ for all $x \in \mathbb{F}^n$.

b) We also know that $Ax_0 \in W$ is the closest vector to y if and only if$y - Ax_0 \in W^+$.

   Therefore, for any $x \in \mathbb{F}^n$ we have that $\langle Ax, y - Ax_0\rangle = 0$. Equivalently, $\langle x, A^*(y - Ax_0)\rangle = 0$ for any x. Thus, $A^*(y - Ax_0) = 0$, which means that $A^*y = A^*Ax_0$.

c) If rank $A = n$, then Lemma 3.2 shows that rank $A^*A = n$.

   Therefore, A^*A is invertible.

   Hence, $x_0 = (A^*A)^{-1}A^*y$.

□

Remark 3.4. Note that a) shows that $A^*Ax_0 = A^*y$ has a solution x_0 for any y.

Therefore, $\text{Im}(A^*A) = \text{Im}\, A^*$.

Suppose now some data points (t_i, y_i) are given for $i \in [1, n] \cap \mathbb{N}$.

Suppose also some line is drawn in the plane of the scatterplot representing the data.

One way to measure how well the line $y = ax + b$ approximates the data is to compute the sum of the squares of the differences between the data points and the corresponding values given by drawing perpendiculars to the line:

$$\delta = \sum_{i=1}^{n} (y_i - at_i - b)^2$$

This method is called the least square approximation.

Let $A = \begin{pmatrix} t_1 & y_1 \\ t_2 & y_2 \\ \vdots & \vdots \\ t_n & y_n \end{pmatrix}$, $x = \binom{a}{b}$ and $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

If $t_1, \ldots, t_n$ contain at least 2 different values, rank $A \geq 2$, so Theorem 3.3 implies there exists the unique best approximation.

4 Normal and Self-Adjoint Operators

Let V be a finite-dimensional inner product space.

Question. When does $T \in \mathrm{Hom}(V, V)$ have an orthonormal basis of eigenvectors?

Answer. If V has such a basis β, then $[T]_\beta = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ for $\lambda_i \in \mathbb{F}$.

By Theorem 2.5,

$\begin{pmatrix} \overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n} \end{pmatrix}$, since β is orthonormal.

Therefore, $[TT^*]_\beta = [T^*T]_\beta$, and thus $TT^* = T^*T$

Definition 4.1. $T \in \mathrm{Hom}(V, V)$ is **normal** if $TT^* = T^*T$ and **self-adjoint** if $T^* = T$.

Similarly, $A \in M_{n\times n}(\mathbb{F})$ can be defined to be normal or self-adjoint.

Theorem 4.2

If $\mathbb{F} = \mathbb{C}$, V has an orthonormal basis of eigenvectors if and only if T is normal.

If $\mathbb{F} = \mathbb{R}$, V has an orthonormal basis of eigenvectors if and only if T is self-adjoint.