MATHEMATICAL TOOLS - PROBLEM SET 8

Due Sunday, January 8th, 23:55, either in the course mailbox or through the Moodle.

Inner Product Spaces. Let $(V, \langle\cdot,\cdot\rangle_V)$ and $(W, \langle\cdot,\cdot\rangle_W)$ be finite dimensional inner-product spaces (over $\mathbb{R}$).

Problem 1. Let $T : V \to W$ be a linear map.

(1) Show that there exists a unique linear map $T^* : W \to V$ s.t. for all $v \in V, w \in W$, $\langle Tv, w\rangle_W = \langle v, T^*w\rangle_V$.
(2) Assume $m \geq n$. Show that there exist orthonormal bases $v_1, \ldots, v_n$ for V and $w_1, \ldots, w_m$ for W and $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$ s.t. for all $v \in V$, $Tv = \sum_{i=1}^{n} \langle v, v_i\rangle_V \sigma_i w_i$ and for all $w \in W$, $T^*w = \sum_{i=1}^{n} \langle w, w_i\rangle_W \sigma_i v_i$.

Part 2 is a formulation of the SVD theorem for arbitrary finite-dimensional inner-product spaces. You may use the (regular) SVD theorem in your answer.

Problem 2. In this problem we'll establish a one-to-one correspondence between inner products on V and positive-definite matrices. Recall that a symmetric matrix $A \in M_n(\mathbb{R})$ is *positive-definite* if for all $x \in \mathbb{R}^n$, $\langle Ax, x\rangle \geq 0$ with equality iff $x = 0$.

Let $v_1, \ldots, v_n$ be an orthonormal basis for V, according to the inner-product $\langle\cdot,\cdot\rangle_V$ (in other words, for all $i, j \in [n]$, $\langle v_i, v_j\rangle_V = \delta_{i,j}$).

(1) Let $\langle\cdot,\cdot\rangle_1$ be an inner product on V. Show that there exists a unique matrix $A \in M_n(\mathbb{R})$ s.t. for all $u, v \in V$, $\langle u, v\rangle_1 = \sum_{i=1}^{n}\sum_{j=1}^{n} \langle u, v_i\rangle_V \langle v, v_j\rangle_V A_{i,j}$.
(2) Show that A is symmetric and positive definite.
(3) Let B be symmetric and positive definite. Define the linear map $T_B : V \to V$ by $T_B(v_i) = \sum_{j=1}^{n} B_{i,j} v_j$. Show that $\langle x, y\rangle_B = \langle T_B x, y\rangle_V$ is an inner product on V.

Positive Semi-Definite Matrices.

Definition. A symmetric matrix $A \in M_n(\mathbb{R})$ is called *positive semi-definite* if for all $x \in \mathbb{R}^n$, $x^T A x \geq 0$.

Definition. $\emptyset \neq C \subseteq \mathbb{R}^m$ is a *cone* if for all $x, y \in C$ and $\alpha, \beta \geq 0$, $\alpha x + \beta y \in C$. In other words, a non-empty set is a cone if it contains all of its non-negative linear combinations.

Problem 3. Let $P_n = \{A \in M_n(\mathbb{R}) : A\, is\, symmetric\, and\, positive\, semi-definite\}$.

(1) Show that P_n is a cone.
(2) Let $v \in \mathbb{R}^n$. Show that $vv^T \in P_n$.
(3) Let $A \in P_n$. Show that there exist orthogonal $v_1, \ldots, v_n \in \mathbb{R}^n$ and $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$ s.t. $A = \sum_{i=1}^{n} \lambda_i v_i v_i^T$.

Variational Characterization of Singular Values. Recall the definitions from recitation: Let $A \in M_{m,n}(\mathbb{R})$ with $m \geq n$. Unit vectors $v \in \mathbb{R}^n, u \in \mathbb{R}^m$ are called (respectively) right-singular and left-singular with singular value $\sigma \geq 0$ if $Av = \sigma u$ and $A^T u = \sigma v$.

Problem 4. Let $\lambda \in \mathbb{R}$ be an eigenvalue of $A^T A$.

(1) Show that $\lambda \geq 0$.
(2) Show that there exist a right-singular vector $v \in \mathbb{R}^n$ and a left-singular vector $u \in \mathbb{R}^m$ with singular value $\sqrt{\lambda}$.
(3) Show that if $\sigma \geq 0$ is a singular value of A with singular vectors v, u as above, then σ^2 is an eigenvalue of $A^T A$ and AA^T with eigenvectors v and u, respectively.

We'll use the following notation: For $k \geq 1$, $B^k = \left\{x \in \mathbb{R}^k : \|x\|_2 = 1\right\}$ (in other words B^k is the set of unit vectors in $\mathbb{R}^k$)[1]. We proved the following theorem in recitation:

Theorem. *Let $A \in M_{m,n}(\mathbb{R})$. Let $v \in B^n, u \in B^m$ and let $\sigma \geq 0$. The following are equivalent:*

- *v, u are right and left-singular, respectively, with singular value σ.*
- *$f(x, y) = \frac{y^T A x}{\|x\|_2 \|y\|_2}$ (defined on $(\mathbb{R}^n \setminus \{0\}) \times (\mathbb{R}^m \setminus \{0\})$) has a critical point at (v, u) and $f(v, u) = \sigma$.*

The next question uses this characterization to prove the SVD theorem, for which we state the following formulation:

Theorem. *(SVD theorem) Let $A \in M_{m,n}(\mathbb{R})$, and assume $m \geq n$. Then there exist orthonormal bases $v_1, \ldots, v_n$ for $\mathbb{R}^n$ and $w_1, \ldots, w_m$ for $\mathbb{R}^m$ and $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$ s.t. for all $1 \leq i \leq n$, v_i and u_i are right and left-singular, respectively, with singular value σ_i.*

Problem 5. We'll prove the theorem by induction on n.

(1) Show that if $n = 1$ the theorem holds.
(2) Show that f attains a **global** maximum somewhere on $B^n \times B^m$. In other words, show that there exist $(v, u) \in B^n \times B^m$ s.t. for all $(x, y) \in (\mathbb{R}^n \setminus \{0\}) \times (\mathbb{R}^m \setminus \{0\})$, $f(v, u) \geq f(x, y)$.
(3) Explain why $f(v, u)$ is a critical point of f and show that $f(v, u) \geq 0$.
The theorem from recitation now tells us that if we set $v_1 = v, u_1 = u, \sigma_1 = f(v, u)$ then v_1, u_1, σ_1 are singular.
(4) Show that if $x \in \{v_1\}^\perp$ then $Ax \in \{u_1\}^\perp$. Thus the linear map $T : \{v_1\}^\perp \to \{u_1\}^\perp$ defined by $Tx = Ax$ is well-defined.
(5) Note that $\dim \{v_1\}^\perp = n - 1, \dim \{u_1\}^\perp = m - 1$. Let $x_1, \ldots, x_{n-1}$ and $y_1, \ldots, y_{m-1}$ be orthonormal bases for $\{v_1\}^\perp$ and $\{u_1\}^\perp$, respectively. Let $\tilde{A} \in M_{m-1,n-1}(\mathbb{R})$ be the matrix representing T according to these bases. Apply the inductive hypothesis w.r.t. $\tilde{A}$ and conclude the result.

[1] A.k.a. the $(k-1)$-dimensional sphere S^{k-1}.