

Problem 1 (10 pts)

Consider a tensor $T = \sum_{i=1}^r u_i \otimes v_i \otimes w_i \in \mathbb{R}^{d_1 \times d_2 \times d_3}$. Recall that $T_a := \sum_{i=1}^r (w_i)_a u_i \otimes v_i$. Prove that for all $a \in [d_3]$, we have

- $\text{colspan}(T_a) \subseteq \text{span}(\{u_i\})$
- $\text{rowspan}(T_a) \subseteq \text{span}(\{v_i\})$

Problem 2 (5 pts)

Show that for a matrix M , its rank and its border rank are always the same. In particular, suppose you are given a matrix M and a parameter r so that for every $\epsilon > 0$ there is a rank r matrix M_r so that M and M_r are entrywise ϵ -close. Show that $\text{rank}(M) \leq r$.

Hint: Use the Eckhart-Young Theorem.

Problem 3 (45 pts)

Let $T = \sum_{i=1}^d \lambda_i u_i^{\otimes 3}$ for $u_1, \dots, u_d \in \mathbb{R}^d$ a collection of orthonormal vectors and $\lambda_1, \dots, \lambda_d > 0$. Let $p : \mathbb{R}^d \rightarrow \mathbb{R}$ be the associated cubic polynomial

$$p(x) := \sum_{i=1}^d \lambda_i \langle u_i, x \rangle^3.$$

The result we will prove is the following.

Theorem 1. *Let $x \in \mathbb{S}^{d-1}$ be any point for which $p(x) \geq 0$. Then x is a strict local maximum¹ for p over the domain \mathbb{S}^{d-1} if and only if $x = u_i$.*

In other words, there are no “spurious local maxima” for the objective function p , which means that we can find a component u_i of the tensor T simply by running an appropriate implementation of gradient descent.

1. **(2 pts)** It suffices to establish [Theorem 1](#) when u_1, \dots, u_d are the standard basis vectors, i.e. when $u_i = (0, \dots, 0, 1, 0, \dots, 0)$ where the 1 entry is in the i -th coordinate. Give an informal argument for why this is the case.

For the remaining sub-problems, we assume that $\{u_i\}$ are the standard basis vectors.

2. **(10 pts)** Prove that every u_i is a strict local maximum.

¹We say a point $x \in \mathbb{S}^{d-1}$ is a strict local maximum for a function $f : \mathbb{S}^{d-1} \rightarrow \mathbb{R}$ if there exists some $\epsilon > 0$ such that for all $x' \in \mathbb{S}^{d-1}$ which are distinct from x and which satisfy $\|x' - x\| \leq \epsilon$, we have $f(x') < f(x)$.

3. **(2 pts)** Prove that any local maximum x of p over \mathbb{S}^{d-1} is a *stationary point*, that is, the projection of the gradient of p at x to the tangent space at x is zero.
4. **(3 pts)** Prove that any local maximum x satisfies

$$\lambda_i x_i^2 = p(x) \cdot x_i \quad \text{for all } i \in [d]. \quad (1)$$

(Here x_i denotes the i -th coordinate of x .)

5. **(3 pts)** Prove that if $x \in \mathbb{S}^{d-1}$ satisfies $p(x) = 0$, then x is *not* a strict local maximum.

Now we arrive at the trickiest part of the proof. Let $x \in \mathbb{S}^{d-1}$ be any point that satisfies $p(x) > 0$ and has more than one nonzero coordinate. We need to show that such an x cannot be a local maximum. Let $S^* \subseteq [d]$ denote the subset of coordinates of x which are nonzero.

6. **(5 pts)** Let $S \subseteq S^*$ be any proper subset of S^* . Construct a unit vector w such that (1) $\langle w, x \rangle = 0$, (2) the entries of w indexed by S are given by scaling all the entries of x indexed by S by the same factor, and (3) the entries of w indexed by $[d] \setminus S$ are given by scaling all the entries of x indexed by $[d] \setminus S$ by the same factor.
7. **(20 pts)** Prove that x is not a strict local maximum.
Hint: Perturb x in the direction of w and rescale it to give a unit vector $x' = \delta w + \sqrt{1 - \delta^2} x$. Argue that p achieves a higher value at x' than at x . You may find [Eq. \(1\)](#) helpful.