MathTools HW 6

1. **Monotonicity of p norms.** Prove that $\|x\|_p$ is decreasing in p. That is, if $1 \le p \le q \le \infty$ then for every x, $\|x\|_p \ge \|x\|_q$.

   Hint: It suffices to show this assuming $\|x\|_p = 1$. Explain why.

2. **Some matrix norms and singular values.** We say that a matrix norm $\|\cdot\|$ is *orthogonally invariant* if for all $A \in M_{m\times n}(\mathbb{R})$, and any $U \in O(m)$ and $V \in O(n)$ one has $\|A\| = \|UAV\|$.[1]

   (a) Show that the Frobenius norm and the ℓ_2-to-ℓ_2 operator norm are orthogonally invariant.

   (b) Deduce that

      i. $\|A\|_{2,2} = \sigma_1(A)$

      ii. $\|A\|_F^2 = \sum_{i=1}^k \sigma_i(A)^2$ (where $k = \min(n, m)$).

   Here $\sigma_1(A) \ge \ldots \ge \sigma_k(A)$ are the singular values of A.

3. **Some operator norms.** Recall that $\|A\|_{p,q}$ is the ℓ_p-to-ℓ_q operator norm of A. Prove the following:

   (a)
   $$\|A\|_{\infty,\infty} = \max_{1\le i\le m} \sum_{j=1}^n |A_{ij}|\,,$$
   the maximum ℓ_1 norm of a row of A.

   (b)
   $$\|A\|_{1,1} = \max_{1\le j\le n} \sum_{i=1}^m |A_{ij}|\,,$$
   the maximum ℓ_1 norm of a column of A.

   (c)
   $$\|A\|_{1,\infty} = \max_{1\le i\le m, 1\le j\le n} |A_{ij}|$$
   the maximum entry (in abs. value) of A.

[1]Notation: $O(m)$ is the set of n-by-n orthogonal matrices, that is, matrices U such that $U^\top U = I$.

4. **Vectors with small ℓ_p norm $(p < 2)$ are compressible (approximately sparse).** Prove the following theorem from the recitation: for any $p < 2$, there is a number $C_p > 0$ such that for any $x \in \mathbb{R}^n$,

$$\|x - x_s^*\|_2 \leq \frac{C_p}{s^{1/p-1/2}} \|x\|_p \,,$$

where x_s^* is the best s-sparse approximation of x (I emphasize again: C_p doesn't depend on n).

Follow these steps:

(a) Prove that for all $t > 0$,

$$|\{i \in [n] \,:\, |x_i| \geq t\}| \leq \frac{\|x\|_p^p}{t^p} \,.$$

In words: x has at most $\|x\|_p^p / t^p$ coordinates such that $|x_i| \geq t$.

(b) We may obviously assume without loss of generality that $x_i \geq 0$ for all i, and ordered in decreasing order: $x_1 \geq x_2 \geq \ldots \geq x_n \geq 0$. Deduce that $x_k \leq k^{-1/p} \|x\|_p$.

(c) Deduce the theorem.

Hint: You may find the following useful: if $f(\cdot)$ is decreasing, then $f(k) \leq \int_{k-1}^{k} f(x)dx$.

5. **The Welch bound.** In many cases we want to construct a set of m vectors $x_1, \ldots, x_m \in S^{n-1}$ (recall that S^{n-1} denotes the Euclidean unit sphere in $\mathbb{R}^n$) such that their *mutual coherence* $\max_{i \neq j} |\langle x_i, x_j \rangle|$ is as small as possible. Obviously, when $m \leq n$, we can take the x_i-s to be orthonormal, giving optimal coherence 0. The interesting case is when $m > n$. You will prove the following lower bound:

$$\max_{i \neq j} |\langle x_i, x_j \rangle| \geq \sqrt{\frac{m-n}{n(m-1)}} \,.$$

Do this in two steps.

(a) Define the m-by-m symmetric matrix $G_{ij} = \langle x_i, x_j \rangle$. Prove that

$$\|G\|_F^2 \geq \frac{m^2}{n} \,.$$

Hint: You may follow this sequence:

i. Explain why G is *positive semi-definite*, meaning all its eigenvalues are non-negatives $\lambda_i \geq 0$.

ii. Prove that

$$\|G\|_F^2 \geq \frac{(\mathrm{tr}(G))^2}{\mathrm{rank}(G)} \,.$$

iii. Use EX4 Q1(c) to bound the rank.

(b) By upper bounding $\|G\|_F^2$, deduce the Welch bound.

Remarks: This is used, for example, in multi-user communications, where the vectors x_i encode messages, and the correlation $|\langle x_i, x_j \rangle|$ corresponds to an *interference* between two users. Another applications is in signal processing and compressed sensing, where it is known that matrices whose columns have small coherence constitute good sensing matrices.