ELEG5481 Signal Processing Optimization Techniques
Tutorial Solution 1

Jan. 24, 2013

Q1. Let x and y be two vector in $\mathbf{R}^n$, and $y \neq 0$. Show that x can be decomposed uniquely in the form of $x = x_\perp + x_\|$, where $x_\| = cy$ for some c, and $x_\perp^T y = 0$. Moreover, $\|x_\|\|_2^2 + \|x_\perp\|_2^2 = \|x\|_2^2$.

Solution: $(x - cy)^T y = 0$ implies that $c = \frac{x^T y}{\|y\|_2^2}$. Other things are easy.

Q2. Prove the following inequalities:

(a) Cauchy-Schwartz inequality:

$$|x^T y| \leq \|x\|_2 \|y\|_2,$$

where the equality holds if and only if $x = cy$ for some c.

(b) Hölder inequality:

$$|x^T y| \leq \|x\|_p \|y\|_q,$$

where $1/p + 1/q = 1$, $p \geq 1$ and $q \geq 1$.

Solution:

(a) Assume that $y \neq 0$; for otherwise, the inequality holds trivially. By **Q1**, we have

$$|x^T y| = |(x_\perp + x_\|)^T y| = |x_\|^T y| = |c| \|y\|_2^2 = \|x_\|\|_2 \|y\| \leq \|x\|_2 \|y\|_2,$$

where $c = \frac{x^T y}{\|y\|_2^2}$, $x_\| = cy$, $x_\perp = x - x_\|$. The equality holds if and only if $\|x_\|\| = \|x\|$. This is the same as $\|x_\perp\| = 0$, which is equivalent to $x_\perp = 0$ by the property of norm. Therefore, equality holds if and only if $x = x_\| = cy$.

(b) The proof is taken from Stephen Boyd's textbook. Assume neither x nor y is zero; for otherwise, the inequality holds true trivially. We will need the following inequality:

$$a^\theta b^{1-\theta} \leq \theta a + (1 - \theta) b,$$

where $a, b \geq 0$ and $0 \leq \theta \leq 1$. By setting

$$a = \frac{|x_i|^p}{\sum_{j=1}^n |x_j|^p}, \; b = \frac{|y_i|^q}{\sum_{j=1}^n |y_j|^q}, \; \theta = 1/p,$$

yields

$$\left(\frac{|x_i|^p}{\sum_{j=1}^n |x_j|^p} \right)^{1/p} \left(\frac{|y_i|^q}{\sum_{j=1}^n |y_j|^q} \right)^{1/q} \leq \frac{|x_i|^p}{p \sum_{j=1}^n |x_j|^p} + \frac{|y_i|^q}{q \sum_{j=1}^n |y_j|^q}.$$

Summing over $i = 1, \ldots, n$, we have

$$\left(\frac{\sum_{i=1}^n |x_i y_i|}{(\sum_{j=1}^n |x_j|^p)^{1/p} (\sum_{j=1}^n |y_j|^q)^{1/q}} \right) \leq \frac{\sum_{i=1}^n |x_i|^p}{p \sum_{j=1}^n |x_j|^p} + \frac{\sum_{i=1}^n |y_i|^q}{q \sum_{j=1}^n |y_j|^q},$$

which is the same as

$$\sum_{i=1}^n |x_i y_i| \leq \left(\frac{1}{p} + \frac{1}{q} \right) \|x\|_p \|y\|_q$$

Q3. Prove the following functions are norms:
(a) $f(x) = \|x\|_\infty$.

(b) $f(z) = \|z\|_* = \sup\{z^T x \mid \|x\| \leq 1\}$ where $\|\cdot\|$ is a norm on $\mathbb{R}^n$.

Solution:
(a) We only verify the triangular inequality. Others properties are easy. It is true that $f(x+y) = |x_i + y_i| \leq |x_i| + |y_i|$ for some index i. We also have $|x_i| \leq f(x)$ and $|y_i| \leq f(y)$. Hence we have $f(x+y) \leq f(x) + f(y)$.
(b) 1. We show $f(z) = 0$ i.f.f. $z = 0$. We have $f(0) = \sup\{0 \mid \|x\| \leq 1\} = 0$. On the other hand, $f(z) = 0$ implies $z^T x \leq 0$ for all x such that $\|x\| \leq 1$. By setting $x = e_i/\|e_i\|$, where e_i is a zero vector except the ith entry being one, we have $z_i \leq 0$; again by setting $x = -e_i/\|e_i\|$, we have $-z_i \leq 0$. Therefore $z_i = 0$. Since this is true for all i, we conclude that $z = 0$.
2. We show $f(z) \geq 0$. Assume $z \neq 0$. We have $f(z) = \sup\{z^T x \mid \|x\| \leq 1\}$. By setting $x = z/\|z\|$, we have $f(z) \geq z^T z/\|z\| > 0$.
3. We show $f(cz) = |c| f(z)$. This is obviously true for $c = 0$. For $c > 0$, we have $f(cz) = \sup\{cz^T x \mid \|x\| \leq 1\} = c \sup\{z^T x \mid \|x\| \leq 1\} = cf(z)$. For $c < 0$, we have $f(cz) = \sup\{cz^T x \mid \|x\| \leq 1\} = \sup\{|c| z^T(-x) \mid \|-x\| \leq 1\} = |c| \sup\{z^T(-x) \mid \|-x\| \leq 1\} = |c| \sup\{z^T x \mid \|x\| \leq 1\}$.
4. We show $f(z+y) \leq f(z) + f(y)$. We have $f(z+y) = \sup\{(z+y)^T x \mid \|x\| \leq 1\} \leq \sup\{z^T x \mid \|x\| \leq 1\} + \sup\{y^T x \mid \|x\| \leq 1\} = f(z) + f(y)$. Think about why the inequality is true.

Q4. Prove the function $f(x) = \|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ with $0 < p < 1$ is not a norm.

Solution:
The triangular inequality does not hold. For example, $f(e_1 + e_2) = 2^{\frac{1}{p}} > f(e_1) + f(e_2) = 2$. Note that even for $0 < p < 1$ people usually call $\|x\|_p$ p-norm, though it is not a norm. Sometimes, this is confusing.

Q5. Prove the following statements
(a) If $\mathcal{S}$ is a nonempty subspace in $\mathbf{R}^n$, then $\mathcal{S} = \mathcal{R}(A)$ for some matrix A.

(b) For any $A \in \mathbf{R}^{m \times n}$, $\mathcal{N}(A) = \mathcal{R}(A^T)^\perp$, where $\mathcal{R}(A^T)^\perp$ is the orthogonal complement of the range space spanned by A^T.

Solution:
(a) If $\mathcal{S}$ is not empty, then there exists some vector $a_1 \in \mathcal{S}$. Because $\mathcal{S}$ is a subspace, we have $\mathcal{R}([a_1]) \subset \mathcal{S}$. If $\mathcal{R}([a_1]) = \mathcal{S}$, then we are done; otherwise, there exists a vector a_2 such that $a_2 \in \mathcal{S}$ but $a_2 \notin \mathcal{R}([a_1])$. Then we have $\mathcal{R}([a_1, a_2]) \subset \mathcal{S}$. Continuing this process, we will find a set of $k < n$ vectors $\{a_i\}_{i=1}^k$ such that $\mathcal{R}([a_1, \ldots, a_k]) = \mathcal{S}$ or a set of n vectors $\{a_i\}_{i=1}^n$ such that $\mathcal{R}([a_1, \ldots, a_n]) \subset \mathcal{S}$. In the former case, we are done by setting $A = [a_1, \ldots, a_k]$. In the latter case, indeed we are done as well by setting $A = [a_1, \ldots, a_k]$. The reason is that there is no vector a_{n+1} such that $a_{n+1} \in \mathcal{S}$ and $a_{n+1} \notin \mathcal{R}([a_1, \ldots, a_n])$, as the way we construct $\{a_i\}_{i=1}^n$ makes sure that they are linear independent and there are at most n linear independent vectors in $\mathbf{R}^n$.
(b) If $x \in \mathcal{N}(A)$, then $Ax = 0$, implying $x^T A^T y = 0$ for all $y \in \mathbf{R}^n$. By definition, we have $x \in \mathcal{R}(A^T)^\perp$. Conversely, if $x \in \mathcal{R}(A^T)^\perp$, then $x^T A^T y = (Ax)^T y = 0$ for all $y \in \mathbf{R}^n$. Setting

$y = e_i$, we have the ith element of Ax is zero. This is true for all i. So $Ax = 0$.