| Convex Optimization and Modeling | Jun.-Prof. Matthias Hein |
|---|---|
| Exercise Sheet 1 - 14.4.2010 | due: 23.4.2010 |

# Exercise 1 - Dual norm

The algebraic dual $(\mathbb{R}^n)^*$ of the vector space $\mathbb{R}^n$ is the set of all linear maps from $\mathbb{R}^n$ to $\mathbb{R}$. Given that the vector space $\mathbb{R}^n$ is equipped with a norm $\|\cdot\|$ one defines a **dual norm** on the dual space $(\mathbb{R}^n)^*$ as

$$\|v\|^* = \sup_{u\in\mathbb{R}^n} \Big\{ \sum_{i=1}^n v_i u_i \;\Big|\; \|u\| \leq 1 \Big\}.$$

This is basically the operator norm of the linear map, $v : \mathbb{R}^n \to \mathbb{R}$, discussed in the lecture.

a. **(2 Points)** Derive the dual norm of the $l_1$-norm, $\|u\|_1 = \sum_{i=1}^n |u_i|$.

b. **(2 Points)** Derive the dual norm of the $l_2$-norm, $\|u\|_2 = \sqrt{\sum_{i=1}^n u_i^2}$.

c. **(2 Points)** Derive the dual norm of the $l_\infty$-norm, $\|u\|_\infty = \max_{i=1,\ldots,n} |u_i|$.

**Hint:**

- first prove a lower bound for $\|v\|^*$ by plugging in a particular $u$, then prove an upper bound on $\|v\|^*$ and show that upper and lower bound agree,
- for $b$) you may use the **Cauchy-Schwarz inequality**

$$|\langle u, v\rangle| \leq \|u\|_2 \|v\|_2 ,$$

or in coordinates

$$|\sum_{i=1}^n u_i v_i| \leq \sqrt{\sum_{i=1}^n u_i^2}\sqrt{\sum_{i=1}^n v_i^2}.$$

**Solution:**

a. For given $v \in \mathbb{R}^n$ with $v \neq 0$ use $u_i = \begin{cases} 0 & \text{if } |v_i| \neq \max_{i=1,\ldots,n} |v_i| \\ v_i/|v_i| & \text{else .} \end{cases}$, where in the case of a non-unique maximum we set an arbitrary component which achieves the maximum to the sign of the component. Note that $\|u\|_1 = 1$. Plugging this particular $u$ into the definition of the dual norm we get

$$\|v\|^* \geq \sum_{i=1}^n u_i v_i = \max_{i=1,\ldots,n} |v_i|.$$

Moreover, using

$$\sum_{i=1}^n u_i v_i \leq \sum_{i=1}^n |u_i v_i| \leq \sum_{i=1}^n |u_i| \sup_{j=1,\ldots,n} |v_j| \leq \|u\|_1 \max_{j=1,\ldots,n} |v_j|.$$

we have

$$\|v\|^* \leq \max_{j=1,\ldots,n} |v_j|.$$

and thus $\|v\|^* = \|v\|_\infty$.

b. For given $v \in \mathbb{R}^n$ with $v \neq 0$ use $u_i = v_i / \|v\|$. Note that $\|u\|_2 = 1$. Plugging this particular $u$ into the definition of the dual norm we get

$$\|v\|^* \geq \sum_{i=1}^n u_i v_i = \frac{1}{\|v\|} \sum_{i=1}^n v_i^2 = \|v\|_2 .$$

Moreover, using Cauchy-Schwarz yields

$$\sum_{i=1}^n u_i v_i \leq \sqrt{\sum_{i=1}^n u_i^2} \sqrt{\sum_{i=1}^n v_i^2} = \|u\|_2 \|v\|_2 .$$

Thus we have

$$\|v\|^* \leq \|v\|_2 .$$

and thus $\|v\|^* = \|v\|_2$.

c. For given $v \in \mathbb{R}^n$ use $u_i = \begin{cases} \text{sign } v_i = v_i/|v_i| & v_i \neq 0 \\ 0 & v_i = 0. \end{cases}$. Note that $\|u\|_\infty = 1$. Plugging this particular $u$ into the definition of the dual norm we get

$$\|v\|^* \geq \sum_{i=1}^n u_i v_i = \sum_{i=1}^n |v_i|.$$

Moreover, using

$$\sum_{i=1}^n u_i v_i \leq \sum_{i=1}^n |u_i v_i| \leq \max_{j=1,\ldots,n} |u_j| \sum_{i=1}^n |v_i| \leq \|u\|_\infty \sum_{i=1}^n |v_i|.$$

we have

$$\|v\|^* \leq \sum_{i=1}^n |v_i|.$$

and thus $\|v\|^* = \|v\|_1$.

# Exercise 2 - Reminder of Linear Algebra and Analysis

a. **(3 Points)** Proof the assertion from the lecture that every real, symmetric matrix $A$ has the decomposition

$$A = Q \Lambda Q^T,$$

where $Q$ is an orthogonal matrix and $\Lambda$ is a diagonal matrix having the eigenvalues on the diagonal.

b. **(3 Points)** The distance of a point $x$ to a set $C$ is defined as

$$d(x, C) = \inf \left\{ \|x - y\| \;\middle|\; y \in C \right\}.$$

Let $C$ be closed. Prove that the distance $d(x, C)$ is realized by an element of $C$ that means $\exists z \in C$ such that $d(x, C) = d(x, z)$.

**Solution:**

a. The eigenvectors $q_i$ (without loss of generality we assume that $\|q_i\| = 1$) of a symmetric matrix are real,

$$\lambda_i = \lambda_i \|q_i\|^2 = \langle q_i, A q_i \rangle = \left\langle A^T q_i, q_i \right\rangle = \overline{\lambda_i} \|q_i\|^2 = \overline{\lambda_i}.$$

and thus $\lambda_i$ is real.
Moreover, the eigenvectors are orthogonal to each other

$$\lambda_i \langle q_i, q_j \rangle = \langle Aq_i, q_j \rangle = \left\langle q_i, A^T q_j \right\rangle = \langle q_i, Aq_j \rangle = \lambda_j \langle q_i, q_j \rangle .$$

Thus $(\lambda_i - \lambda_j) \langle q_i, q_j \rangle = 0$. Thus, if $\lambda_i \neq \lambda_j$ we have $\langle q_i, q_j \rangle = 0$. If an eigenvalue has a multiplicity larger than 1 we use an orthonormal basis of the resulting eigenspace.

All eigenvectors plus an orthnormal basis of the kernel (or null space) of $A$ thus provide a basis of $\mathbb{R}^n$. Now, it is a standard result in linear algebra that the representation of a matrix $A$ in another basis is given by

$$A = SBS^{-1},$$

where $S$ contains as columns the new basis vectors represented in terms of the old basis and $B_{ij} = \langle q_i, Aq_j \rangle$ are the components with respect to the new basis. In our case $S = Q$ and since $Q$ is an orthogonal matrix we have $Q^{-1} = Q^T$. Moreover, $B_{ij} = \langle q_i, Aq_j \rangle = \lambda_j \delta_{ij}$, where $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$. Thus $A = Q\Lambda Q^T$.

b. The function $y \rightarrow \|x - y\|$ for fixed $x$ is clearly continuous. Pick an arbitrary $w \in C$, then in order to compute the distance $d(x, C)$ it is sufficient to minimize over the set $\{y \in C \mid \|x - y\| \leq \|x - w\|\}$, which is closed and bounded and therefore compact. A continuous function attains its minimum on a compact set and thus there exists a $z \in C$ such that $d(x, z) = d(x, C)$.