Chapter 5

Duality

Contents

5.1

Introduction

Overview

•

Dual spaces are the foundation for

- converting between maximization problems and minimization problems,
- gradients (for optimality conditions),
- Lagrange multipliers (constrained optimization).

5.2

Linear functionals

$f : \mathcal{X} \to \mathcal{F}$ where $\mathcal{X}$ is a vector space over $\mathcal{F} = \mathbb{R}$ or $\mathcal{F} = \mathbb{C}$.
However, Luenberger only considers real vector spaces and hence only real functionals.

When $\mathcal{X}$ is a normed space, we can discuss: continuity of f, boundedness of f, the (operator) norm of f, denoted $|||f|||$.

Example. $\mathcal{X} = (C[0,1], \|\cdot\|_\infty)$, with $f(\boldsymbol{x}) = x(1/3)$. This f is linear. Of course it is bounded too:

$$|f(\boldsymbol{x})| = |x(1/3)| \le \sup_{t\in[0,1]} |x(t)| = \|\boldsymbol{x}\|_\infty .$$

Considering $x(t) = 1$ shows that $|||f||| = 1$.

Example. In contrast, in $\mathcal{X} = (C[0,1], \|\cdot\|_2)$ the same linear functional is unbounded! Because $|f(\boldsymbol{x})| / \|\boldsymbol{x}\|$ can be made arbitrarily large by choosing by choosing $x(t)$ to be a sufficiently thin triangle function centered at $1/3$.

Normed Dual

$B(\mathcal{X}, \mathbb{C})$ is the normed space of bounded linear functionals on a normed space $\mathcal{X}$.

Definition. $B(\mathcal{X}, \mathbb{C})$ is called the **normed dual** of $\mathcal{X}$ and is denoted $\mathcal{X}^\star$.

Theorem. *$\mathcal{X}^\star$ is a Banach space.*

Proof. This follows directly from the fact that $\mathbb{C}$ is complete, and, as shown in Ch. 6, $B(\mathcal{X}, \mathcal{Y})$ is complete if $\mathcal{Y}$ is complete. □

Note that the dual of a normed space is a space of linear functionals, which at first might seem like quite different entities. We will see that often they are not so different.

5.3

Examples of duals

Example. The normed dual of $\mathbb{E}^n$. What is the class of linear functionals on $\mathbb{E}^n$?

Suppose $f : \mathbb{E}^n \to \mathbb{R}$ is linear. Then define $\alpha_i = f(\boldsymbol{e}_i) \in \mathbb{R}$, where $\boldsymbol{e}_i$ is the ith unit vector in $\mathbb{R}^n$.
When $\boldsymbol{x} = (a_1, \ldots, a_n)$ then $f(\boldsymbol{x}) = f(\sum_{i=1}^n a_i \boldsymbol{e}_i) = \sum_{i=1}^n a_i f(\boldsymbol{e}_i) = \sum_{i=1}^n a_i \alpha_i = \langle \boldsymbol{x}, \boldsymbol{\alpha} \rangle$, where $\boldsymbol{\alpha} \in \mathbb{R}^n$.
Furthermore, as shown previously, $|||f||| = \sqrt{\sum_{i=1}^n \alpha_i^2} = \|\boldsymbol{\alpha}\|_2$.
So $(\mathbb{E}^n)^\star = \{f(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{\alpha} \rangle : \boldsymbol{\alpha} \in \mathbb{R}^n\}$, with $|||f||| = \|\boldsymbol{\alpha}\|_2$. Thus the normed dual of $\mathbb{E}^n$ is "essentially" $\mathbb{E}^n$ since f is determined by the coefficients $\boldsymbol{\alpha} \in \mathbb{E}^n$ and the norm of f is the 2-norm of $\boldsymbol{\alpha}$.

To characterize "essentially" more precisely: the normed dual of $\mathbb{E}^n$ is isometrically isomorphic to $\mathbb{E}^n$.

(See ℓ_2 below for essence of proofs.)

Similarly, $(\mathbb{C}^n)^\star = \{f(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{\alpha} \rangle : \boldsymbol{\alpha} \in \mathbb{C}^n\}$, again with $|||f||| = \|\boldsymbol{\alpha}\|_2$.

Example. The dual of ℓ_p, $1 \leq p < \infty$ is (essentially) ℓ_q, where $1/p + 1/q = 1$.

Theorem. *(5.3-1) For $p \in [1, \infty)$ and $1/p + 1/q = 1$:*

$$f \in \ell_p^\star = B(\ell_p, \mathbb{C}) \iff \forall \boldsymbol{x} = (a_1, a_2, \ldots) \in \ell_p,\ f(\boldsymbol{x}) = \sum_{i=1}^{\infty} \alpha_i a_i,\ \text{where } \boldsymbol{y} = \{\alpha_i\} \in \ell_q \text{ with } \alpha_i \in \mathbb{C}.$$

Furthermore, $|||f||| = \|\boldsymbol{y}\|_q$.

Proof.

Recall $\ell_p = \{\boldsymbol{x} = (a_i) : \sum_{i=1}^{\infty} |a_i|^p < \infty,\ a_i \in \mathbb{C}\}$, $p \in [1, \infty)$.

Claim 0. For $p \in [1, \infty)$, $\boldsymbol{x} = (a_i) \in \ell_p \Longrightarrow \boldsymbol{x} = \sum_{i=1}^{\infty} a_i \boldsymbol{e}_i$, where $\{\boldsymbol{e}_i\}$ denotes the usual unit vectors in ℓ_p.
Pf. Clearly $\sum_{j=1}^{n} a_j \boldsymbol{e}_j \in \ell_p$, so

$$\left\| \boldsymbol{x} - \sum_{j=1}^{n} a_j \boldsymbol{e}_j \right\|_p^p = \sum_{i=1}^{\infty} \left| a_i - \left[\sum_{j=1}^{n} a_j \boldsymbol{e}_j \right]_i \right|^p = \sum_{i=n+1}^{\infty} |a_i|^p \to 0 \text{ as } n \to \infty.$$

Note. Claim 0 is *false* for $p = \infty$!

Claim 1. $f \in \ell_p^\star = B(\ell_p, \mathbb{C}) \Longrightarrow f(\boldsymbol{x}) = \sum_{i=1}^{\infty} \alpha_i a_i$ where $\boldsymbol{y} = \{\alpha_i\} \in \ell_q$.

Pf. Let $\alpha_i = f(\boldsymbol{e}_i)$. For any $\boldsymbol{x} \in \ell_p$, by continuity and linearity of f: $f(\boldsymbol{x}) = \sum_{i=1}^{\infty} a_i f(\boldsymbol{e}_i) = \sum_{i=1}^{\infty} \alpha_i a_i$.

Consider first $1 < p < \infty$. For $n \in \mathbb{N}$ define $\boldsymbol{x}_n = \sum_{i=1}^{n} |\alpha_i|^{q/p} \operatorname{sgn}(\alpha_i)\, \boldsymbol{e}_i$, where $\operatorname{sgn}(\alpha) \triangleq \begin{cases} \alpha^*/|\alpha|, & \alpha \neq 0 \\ 0, & \alpha = 0. \end{cases}$
Note $\alpha \operatorname{sgn}(\alpha) = |\alpha|$.
Then $\|\boldsymbol{x}_n\|_p = \left(\sum_{i=1}^{n} |\alpha_i|^q\right)^{1/p}$ and $f(\boldsymbol{x}_n) = \sum_{i=1}^{n} |\alpha_i|^{q/p} \operatorname{sgn}(\alpha_i)\, f(\boldsymbol{e}_i) = \sum_{i=1}^{n} |\alpha_i|^{q/p+1} = \sum_{i=1}^{n} |\alpha_i|^q$.
But since $f \in B(\mathcal{X}, \mathbb{C})$, $|f(\boldsymbol{x}_n)| \leq |||f|||\, \|\boldsymbol{x}_n\|_p$, so $\sum_{i=1}^{n} |\alpha_i|^q \leq |||f||| \left(\sum_{i=1}^{n} |\alpha_i|^q\right)^{1/p}$.
Rearranging: $\left(\sum_{i=1}^{n} |\alpha_i|^q\right)^{1/q} \leq |||f|||$, $\forall n \in \mathbb{N}$. Hence $\boldsymbol{y} = (\alpha_1, \alpha_2, \ldots) \in \ell_q$ and $\|\boldsymbol{y}\|_q \leq |||f|||$.

For $p = 1$ and $q = \infty$ define $\boldsymbol{x}_n = \operatorname{sgn}(\alpha_n)\, \boldsymbol{e}_n$. So $\|\boldsymbol{x}_n\|_p = |\operatorname{sgn}(\alpha_n)| \leq 1$ and $f(\boldsymbol{x}_n) = \alpha_n \operatorname{sgn}(\alpha_n) = |\alpha_n|$.
Thus $|\alpha_n| = f(\boldsymbol{x}_n) \leq |||f|||\, \|\boldsymbol{x}_n\|_p \leq |||f|||$, and hence $\boldsymbol{y} \in \ell_\infty$ and $\|\boldsymbol{y}\|_\infty \leq |||f|||$.

Claim 2. $\boldsymbol{y} = \{\alpha_i\} \in \ell_q \Longrightarrow f_{\boldsymbol{y}}(\boldsymbol{x}) \triangleq \sum_{i=1}^{\infty} \alpha_i a_i \in B(\ell_p, \mathbb{C}) = \ell_p^\star$ and $|||f_{\boldsymbol{y}}||| = \|\boldsymbol{y}\|_q$.

Pf. For $1 < p < \infty$, by the Hölder inequality, $|f_{\boldsymbol{y}}(\boldsymbol{x})| = |\sum_{i=1}^{\infty} \alpha_i a_i| \leq \|\boldsymbol{y}\|_q \|\boldsymbol{x}\|_p$ so $f_{\boldsymbol{y}} \in B(\mathcal{X}, \mathbb{C})$ and $|||f_{\boldsymbol{y}}||| \leq \|\boldsymbol{y}\|_q$.
Again we have $f_{\boldsymbol{y}}(\boldsymbol{e}_i) = \alpha_i$, so since analysis in Claim 1 applies since $f_{\boldsymbol{y}} \in \ell_p^\star$, $\|\boldsymbol{y}\|_q \leq |||f_{\boldsymbol{y}}|||$, so we conclude $|||f_{\boldsymbol{y}}||| = \|\boldsymbol{y}\|_q$.

For $p = 1$ and $q = \infty$, $|f_{\boldsymbol{y}}(\boldsymbol{x})| = |\sum_{i=1}^{\infty} \alpha_i a_i| \leq \sum_{i=1}^{\infty} (\sup_j \alpha_j)|a_i| = \|\boldsymbol{y}\|_\infty \|\boldsymbol{x}\|_1$, so $f_{\boldsymbol{y}} \in B(\mathcal{X}, \mathbb{C})$ with $|||f_{\boldsymbol{y}}||| \leq \|\boldsymbol{y}\|_\infty$.
Again we have $f_{\boldsymbol{y}}(\boldsymbol{e}_i) = \alpha_i$, so by the previous analysis, $\|\boldsymbol{y}\|_\infty \leq |||f_{\boldsymbol{y}}|||$, so we conclude $|||f_{\boldsymbol{y}}||| = \|\boldsymbol{y}\|_\infty$. □

Remark. The above works for $1 \leq p < \infty$. However, the dual of ℓ_∞ is not ℓ_1.

Remark. One can formalize "essentially" above by showing that $\ell_p^\star$ is **isometrically isomorphic** to ℓ_q (homework). Most of the work is in the above theorem proof.

Example. By similar arguments, the dual of $\mathcal{L}_p[0, 1]$ consists of functions of the form

$$f(\boldsymbol{x}) = \int_0^1 x(t) y(t)\, \mathrm{d}t,$$

where $\boldsymbol{y} \in \mathcal{L}_q[0, 1]$ and $1/p + 1/q = 1$, for $1 \leq p < \infty$. Thus the dual of $\mathcal{L}_p$ is "essentially" $\mathcal{L}_q$.

Normed duals in Hilbert spaces

In an inner product space $(\mathcal{X}, \langle \cdot, \cdot \rangle)$, the functional $f_{\boldsymbol{y}}(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$, for fixed $\boldsymbol{y}$, is a continuous linear functional in $\boldsymbol{x}$.
(Linearity and continuity of $f_{\boldsymbol{y}}$ follow directly from the corresponding properties of inner products.)

Is $f_{\boldsymbol{y}}$ bounded? Yes, by Cauchy-Schwarz: $|f_{\boldsymbol{y}}(\boldsymbol{x})| = |\langle \boldsymbol{x}, \boldsymbol{y} \rangle| \leq \|\boldsymbol{x}\| \|\boldsymbol{y}\|$ so $\|\!|f_{\boldsymbol{y}}|\!\| \leq \|\boldsymbol{y}\|$.
Of course $f_{\boldsymbol{y}}(\boldsymbol{y}) = \|\boldsymbol{y}\|^2$ so in fact $\|\!|f_{\boldsymbol{y}}|\!\| = \|\boldsymbol{y}\|$.

Each distinct vector $\boldsymbol{y}$ produces a distinct functional $f_{\boldsymbol{y}} \in \mathcal{X}^\star$. So $\mathcal{X}^\star$ is quite rich with functions!
But are *all* functionals in $\mathcal{X}^\star$ of the above form? In Hilbert spaces at least, the answer is *yes*.

Theorem. *(Riesz-Fréchet) (aka Riesz Representation Theorem [3, p. 345])*
If $g \in \mathcal{H}^\star = B(\mathcal{H}, \mathcal{F})$ where $\mathcal{H}$ is Hilbert space over $\mathcal{F} = \mathbb{R}$ or $\mathcal{F} = \mathbb{C}$, then
- *there exists a unique vector $\boldsymbol{y} \in \mathcal{H}$ such that $g(\boldsymbol{x}) = f_{\boldsymbol{y}}(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$, $\forall \boldsymbol{x} \in \mathcal{H}$,*
- *$\|\!|f_{\boldsymbol{y}}|\!\| = \|\boldsymbol{y}\| = \|\!|g|\!\|$, and*
- *every $\boldsymbol{y} \in \mathcal{H}$ so determines a unique bounded linear functional on $\mathcal{H}$.*

In particular, when $\mathcal{F} = \mathbb{R}$, the mapping $T : \mathcal{H} \to \mathcal{H}^\star$ defined by $T(\boldsymbol{y}) = f_{\boldsymbol{y}}$ is an onto linear isometry, i.e., *$\mathcal{H}$ and $\mathcal{H}^\star$ are isometrically isomorphic.*

Proof.

. . . .

Claim 1. T is onto $\mathcal{H}^\star$, *i.e.*, $g \in \mathcal{H}^\star \Longrightarrow \exists \boldsymbol{y} \in \mathcal{H}$ s.t. $g(\cdot) = f_{\boldsymbol{y}}(\cdot) = \langle \cdot, \boldsymbol{y} \rangle$. Goal: find the $\boldsymbol{y}$.
Given $g \in \mathcal{H}^\star$, define $N \triangleq \{\boldsymbol{x} \in \mathcal{H} : g(\boldsymbol{x}) = 0\} = N(g)$ (nullspace).
By the linearity of g, N is clearly a subspace of $\mathcal{H}$.
Furthermore, N is closed by the continuity of g (since g is bounded), since $\boldsymbol{x}_n \in N$ and $\boldsymbol{x}_n \to \boldsymbol{x} \Longrightarrow g(\boldsymbol{x}) = \lim_{n\to\infty} g(\boldsymbol{x}_n) = 0$.

If $N = \mathcal{H}$, then $g \equiv 0$ and we simply take $\boldsymbol{y} = \mathbf{0}$.
Otherwise, by the theorem for orthogonal complements: $\mathcal{H} = N \oplus N^\perp$. (Here is where we use completeness!)
Since $N \neq \mathcal{H}$, there exists a nonzero vector $\boldsymbol{z} \in N^\perp$. Since $\boldsymbol{z} \notin N$, $g(\boldsymbol{z}) \neq 0$.
Since $N^\perp$ is a subspace, we can normalize $\boldsymbol{z}$ such that $g(\boldsymbol{z}) = 1$.
For $\boldsymbol{x} \in \mathcal{H}$, $g(\boldsymbol{x} - g(\boldsymbol{x})\boldsymbol{z}) = g(\boldsymbol{x}) - g(g(\boldsymbol{x})\boldsymbol{z}) = g(\boldsymbol{x}) - g(\boldsymbol{x})g(\boldsymbol{z}) = 0$, so $\boldsymbol{x} - g(\boldsymbol{x})\boldsymbol{z} \in N$.

Since $\boldsymbol{z} \perp N$, we have $0 = \langle \boldsymbol{x} - g(\boldsymbol{x})\boldsymbol{z}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle - g(\boldsymbol{x}) \|\boldsymbol{z}\|^2$, so $g(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{z} / \|\boldsymbol{z}\|^2 \rangle$.
Thus choosing $\boldsymbol{y} = \boldsymbol{z} / \|\boldsymbol{z}\|^2$ gives the required representation with $g = f_{\boldsymbol{y}}$.

. . . .

Claim 2. T is invertible, *i.e.*, $\boldsymbol{y}$ is unique.

Suppose $g(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{y}' \rangle, \forall \boldsymbol{x} \in \mathcal{H}$ Then $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y}' \rangle$ so $\langle \boldsymbol{x}, \boldsymbol{y} - \boldsymbol{y}' \rangle = 0$, $\forall \boldsymbol{x} \in \mathcal{H}$ which implies $\boldsymbol{y} - \boldsymbol{y}' = \mathbf{0}$ by a simple Lemma that followed the axioms of inner products (consider $\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{y}'$).

. . . .

Claim 3. T is norm preserving, *i.e.*, $\|\!|T(\boldsymbol{y})|\!\| = \|\boldsymbol{y}\|$.
$\|\!|f_{\boldsymbol{y}}|\!\| = \|\boldsymbol{y}\|$ was argued before the theorem statement.

Claim 4. T is linear when $\mathcal{F} = \mathbb{R}$.
$T(\alpha\boldsymbol{y} + \boldsymbol{z}) = f_{\alpha\boldsymbol{y}+\boldsymbol{z}} = \langle \cdot, \alpha\boldsymbol{y} + \boldsymbol{z} \rangle = \alpha^* \langle \cdot, \boldsymbol{y} \rangle + \langle \cdot, \boldsymbol{z} \rangle = \alpha f_{\boldsymbol{y}} + f_{\boldsymbol{z}} = \alpha T(\boldsymbol{y}) + T(\boldsymbol{z})$, where we used the fact $\alpha^* = \alpha$ for a real Hilbert space. □

So in Hilbert spaces, we have characterized completely *all* the continuous linear functionals!

Of course, we already knew this theorem holds for ℓ_2 and $\mathcal{L}_2$. The generalization here is to arbitrary Hilbert spaces.

Working in an inner product space is not a *necessary* condition since we have seen that, even though ℓ_p is not an inner product space for $p \neq 1$, the normed dual of ℓ_p is $\ell_p^\star = \{f(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle_2 : \boldsymbol{y} \in \ell_q\}$.

To understand why the above Theorem assumes completeness, consider the inner product space $\mathcal{X}$ consisting of sequences with finitely many nonzero terms, along with the usual ℓ_2 inner product, and consider the linear functional $f(\boldsymbol{x}) = \sum_{i=1}^\infty a_i / 2^i$. This is bounded, so $f \in \mathcal{X}^\star$, but it cannot be represented by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ for any $\boldsymbol{y} \in \mathcal{X}$. So completeness is an essential component of the above theorem.

Exercise. Are $\mathcal{H}$ and $\mathcal{H}^\star$ unitarily equivalent? **??**