Duality and optimality in stochastic optimization and mathematical finance

Teemu Pennanen
King’s College London

Ari-Pekka Perkkiö
Technische Universität Berlin

Terryfest, Limoges 18-22 May 2015
• Rockafellar, *Conjugate duality and optimization*, 1974, SIAM.
• Biagini, Pennanen and Perkkiö, *Duality and optimality in stochastic optimization and mathematical finance*, 2015.
**Conjugate duality** studies parametric optimization problems

\[
\text{minimize} \quad F(x, u) \quad \text{over} \quad x \in X, \quad (P)
\]

where the parameter \( u \) takes values in a locally convex space \( U \) in separating duality with \( Y \). If \( F \) is convex on \( X \times U \), then

- the optimum value \( \varphi(u) \) is convex on \( U \),
- the associated Lagrangian

\[
L(x, y) = \inf_{u \in U} \{ F(x, u) - \langle u, y \rangle \}.
\]

is convex concave on \( X \times Y \),
- the conjugate of \( \varphi \) can be expressed as

\[
\varphi^*(y) = \sup_{u \in U} \{ \langle u, y \rangle - \varphi(u) \} = -\inf_{x \in X} L(x, y).
\]
Conjugate duality

- If $\varphi$ is closed, then there is no duality gap: $\varphi = \varphi^{**}$.
- If $\varphi$ is subdifferentiable at $u$, then $x \in X$ solves (P) if and only if $(x, y)$ is a saddle point of $L - \langle \cdot, y \rangle$ for some $y \in Y$.

Note that we have not assumed that $X$ is locally convex.

- In stochastic optimization, this will allow us to choose a large enough $X$ so that primal solutions exist and $\varphi$ is closed.
- On the other hand, we will have to work a bit harder to find explicit expressions for $\varphi^*$ and the saddle point conditions.
Let \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, P)\) be a filtered probability space and consider the parametric optimization problem

\[
\text{minimize } Ef(x, u) := \int f(x(\omega), u(\omega), \omega) dP(\omega) \text{ over } x \in \mathcal{N},
\]

- \(\mathcal{N} = \{(x_t)_{t=0}^T | x_t \in L^0(\Omega, \mathcal{F}_t, P; \mathbb{R}^d)\}\),
- \(u \in L^p(\Omega, \mathcal{F}, P; \mathbb{R}^m)\) is a fixed parameter,
- \(f : \mathbb{R}^{d(T+1)} \times \mathbb{R}^m \times \Omega \to \overline{\mathbb{R}}\) is a convex normal integrand,
- the integral is defined as \(+\infty\) unless the positive part of the integrand is integrable.
Example 1 (Inequality constraints) If

\[ f(x, u, \omega) = \begin{cases} f_0(x, \omega) & \text{if } f_j(x, \omega) + u_j \leq 0 \text{ for } j = 1, \ldots, m, \\ +\infty & \text{otherwise}, \end{cases} \]

where \( f_j \) are convex normal integrands, then \( \varphi(u) \) is the optimal value of the problem

\[
\begin{align*}
\min_{x \in \mathcal{N}} & \quad E f_0(x(\omega), \omega) \\
\text{subject to } & \quad f_j(x(\omega), \omega) + u_j(\omega) \leq 0 \quad j = 1, \ldots, m.
\end{align*}
\]

This was studied by [Rockafellar and Wets, 1978] in the case of bounded strategies.
Stochastic optimization

**Example 2 (Shadow price of information)** If $m = d(T + 1)$ and

$$f(x, u, \omega) = h(x + u, \omega),$$

where $h$ is a convex normal integrand, then the problem becomes the **nonadapted perturbation**

$$\min_{x \in \mathcal{N}} Eh(x + u).$$

of the stochastic optimization problem

$$\min_{x \in \mathcal{N}} Eh(x).$$

This was studied in [Rockafellar and Wets, 1976].
Stochastic optimization

Example 3 (Optimal stopping)

If $d = m = 1$ and

$$f(x, u, \omega) = \begin{cases} 
\sum_{t=0}^{T} x_t Z_t(\omega) & \text{if } x \geq 0 \text{ and } \sum_{t=0}^{T} x_t \leq u, \\
+\infty & \text{otherwise,}
\end{cases}$$

for an adapted real-valued process $Z$, the problem becomes

$$\minimize_{x \in \mathbb{N}_+} \mathbb{E} \sum_{t=0}^{T} x_t Z_t \quad \text{subject to} \quad \sum_{t=0}^{T} x_t \leq u \ P\text{-a.s.}$$

When $u = 1$, this is a convex relaxation of the optimal stopping problem. The relaxation does not affect the optimal value.
Example 4 (Optimal investment) \textit{Let} m = 1 \textit{and}

\[ f(x, u, \omega) = v \left( u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}(\omega) \right) \]

\textit{where} s \textit{is an adapted price process and} v : \mathbb{R} \to \overline{\mathbb{R}} \textit{is convex. The problem becomes}

\[ \min_{x \in \mathcal{N}} E v \left( u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1} \right) \]

\textit{which is the problem of optimal investment with liability} u.
Example 5 (Optimal investment in illiquid markets) Let

\[ f(x, u, \omega) = \begin{cases} 
\sum_{t=0}^{T} v_t(S_t(\Delta x_t, \omega) + u_t) & \text{if } x_t \in D_t(\omega), \ x_T = 0 \\
+\infty & \text{otherwise}
\end{cases} \]

where

- \( S_t : \mathbb{R}^d \times \Omega \to \mathbb{R} \) is such that \( S_t(\cdot, \omega) \) are convex with \( S_t(0, \omega) = 0 \) and \( S_t(x, \cdot) \) are \( \mathcal{F}_t \)-measurable,
- \( \omega \mapsto D_t(\omega) \) is \( \mathcal{F}_t \)-measurable with \( D_t(\omega) \) closed convex and \( 0 \in D_t(\omega) \),
- \( v_t : \mathbb{R} \to \mathbb{R} \) is convex.

The problem becomes

\[
\min_{x \in \mathcal{N}_D} \mathbb{E} \sum_{t=0}^{T} v_t(S_t(\Delta x_t) + u_t).
\]
We will apply conjugate duality with

\[ F(x, u) = \mathbb{E}f(x, u), \]
\[ X = \mathcal{N}, \]
\[ U = L^p(\Omega, \mathcal{F}, P; \mathbb{R}^m), \]
\[ Y = L^q(\Omega, \mathcal{F}, P; \mathbb{R}^m), \]
\[ \langle u, y \rangle = \mathbb{E}(u \cdot y). \]

- Our first aim is to show that, one often has

\[ \varphi^*(y) = - \inf_{x \in \mathcal{N}} L(x, y) = - \inf_{x \in \mathcal{N}^\infty} L(x, y) \]

where \( \mathcal{N}^\infty := \mathcal{N} \cap L^\infty. \)
- This yields explicit expressions for \( \varphi^* \) in many situations.
Duality

Let

\[ \tilde{\varphi}(u) = \inf_{x \in \mathcal{N}^\infty} E f(x, u), \]

\[ \mathcal{N}^\perp = \{ v \in L^1(\Omega, \mathcal{F}, P; \mathbb{R}^n) \mid E(x \cdot v) = 0 \ \forall x \in \mathcal{N}^\infty \}, \]

\[ l(x, y, \omega) = \inf_{u \in \mathbb{R}^m} \{ f(x, u, \omega) - u \cdot y \}. \]

**Theorem 1**  If \( \text{dom } El(\cdot, y) \cap \mathcal{N}^\infty \subseteq \text{dom } \tilde{\varphi} \), then

\[ \tilde{\varphi}^*(y) = -\inf_{x \in \mathcal{N}^\infty} El(x, y). \]

If in addition, there exists \( v \in \mathcal{N}^\perp \) with \( \tilde{\varphi}^*(y) = Ef^*(v, y) \), then

\[ \varphi^*(y) = \tilde{\varphi}^*(y). \]

**Lemma 2 (Perkkiö, 2014)**  If \( x \in \mathcal{N} \), \( v \in \mathcal{N}^\perp \) and \( E[ x \cdot v ]^+ \in L^1 \), then \( E(x \cdot v) = 0 \).
Example 6 (Inequality constraints) In the model

\[
\begin{align*}
\text{minimize} \quad & Ef_0(x(\omega), \omega) \\
\text{subject to} \quad & f_j(x(\omega), \omega) + u_j(\omega) \leq 0 \quad j = 1, \ldots, m,
\end{align*}
\]

the conditions of Theorem 1 hold provided \( f_j(x, \omega) \in L^p \) for all \( x \in \mathbb{R}^n \). This follows from the Mackey-continuity of \( El(\cdot, y) \) on \( L^\infty \), which in turn follows from Rockafellar, Integrals which are convex functionals II, 1971.

We have

\[
l(x, y, \omega) = \begin{cases} 
  f_0(x, \omega) + \sum_{j=1}^{m} y_j f_j(x, \omega) & \text{if } y \geq 0, \\
  -\infty & \text{otherwise}
\end{cases}
\]

but, in general, don’t have explicit expressions for \( \varphi^*(y) \).
Example 7 (Shadow price of information) When
\( f(x, u, \omega) = h(x + u, \omega) \), the conditions of Theorem 1 hold as soon as \( Eh \cap L^p \). We get

\[
\varphi^*(y) = \begin{cases} 
Eh^*(y) & \text{if } E_t y_t = 0 \ \forall t, \\
\infty & \text{otherwise}
\end{cases}
\]

and, in particular,

\[
\varphi^{**}(0) = \sup_{y \in \mathcal{N}^\perp} E[-h^*(y)] = \sup_{y \in \mathcal{N}^\perp} E \inf_{x \in \mathbb{R}^n} \{h(x) - x \cdot y\},
\]

hence the name shadow price of information; see [Rockafellar and Wets, 1976], [Back and Pliska, 1987] and [Davis, 1992]. This allows for MC much as in [Rogers, 2002].
Example 8 (Optimal stopping)  The optimal value of the optimal stopping problem

\[
\text{maximize } \underset{x \in \mathbb{N}_+}{E} \sum_{t=0}^{T} x_t Z_t \quad \text{subject to } \sum_{t=0}^{T} x_t \leq 1
\]

equals that of

\[
\text{minimize } y_0 \quad \text{subject to } y \geq Z^+, \quad y \in \mathcal{M}^\infty
\]

where $\mathcal{M}^\infty$ is the set of bounded martingales and $Z^+$ is the positive part of $Z$. 
Example 9 (Optimal investment) Assume that, \( \forall x \in \mathcal{N}^\infty \),

\[ \exists u \in L^p \text{ such that } EV(u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}) < \infty, \text{ then} \]

\[ \varphi^{**}(u) = \sup_{y \in \mathcal{Q}} E[uy - V^*(y)], \]

where \( \mathcal{Q} \) is the set of positive multiples of martingale densities \( y \in \mathcal{Y}, \) i.e. densities \( dQ/dP \) of probability measures \( Q \ll P \) under which the price process \( s \) is a martingale.
Closedness criteria

- The above expressions for $\varphi^{**}$ provide dual representations of the optimal value $\varphi$ provided $\varphi$ is proper and lower semicontinuous (lsc), i.e.

$$\liminf_{\nu \to \infty} \varphi(u^{\nu}) \geq \varphi(u)$$

whenever $u^{\nu} \to u$ in $L^p$.

- The traditional “direct method” assumes that $Ef$ is jointly lsc and $Ef(\cdot, u)$ is inf-compact uniformly in $u$.

- In financial models, the topological inf-compactness condition often fails but there is a more general measure theoretic counterpart that works well in $\mathcal{N}$.
**Closedness criteria**

**Theorem 3 (Komlós)** If \((x^\nu)_{\nu=1}^\infty \subset L^0(\Omega, \mathcal{F}, P; \mathbb{R}^n)\) is almost surely bounded in the sense that

\[
\sup_{\nu} |x^\nu(\omega)| < \infty \quad P\text{-a.s.}
\]

then there is a sequence of convex combinations \(\bar{x}^\nu \in \text{co}\{x^\mu \mid \mu \geq \nu\}\) that converges almost surely in \(L^0\).

This yields the following infinite-dimensional version of Theorem 8.4 from *Convex Analysis*.

**Theorem 4** Let \(C : \Omega \rightrightarrows \mathbb{R}^n\) be closed convex-valued and \(\mathcal{F}\)-measurable. If \(\{x \in \mathcal{N} \mid x \in C^\infty \text{ a.s.}\} = \{0\}\), then every sequence in \(\{x \in \mathcal{N} \mid x \in C \text{ a.s.}\}\) is almost surely bounded.
Theorem 5  Assume that $f$ is bounded from below and that

$$\{x \in \mathcal{N} | f^\infty(x(\omega), 0, \omega) \leq 0 \text{ a.s.}\}$$

is a linear space. Then

$$\varphi(u) = \inf_{x \in \mathcal{N}} E f(x, u)$$

is closed on $L^p$ and the inf is attained for every $u \in L^p$.

The lower bound has been relaxed in [Perkkiö, 2014].
Example 10 (Optimal stopping) When

\[
f(x, u, \omega) = \begin{cases} 
- \sum_{t=0}^{T} x_t Z_t(\omega) & \text{if } x \geq 0 \text{ and } \sum_{t=0}^{T} x_t \leq u, \\
+\infty & \text{otherwise},
\end{cases}
\]

we have \(f^\infty = f\) and

\[
\{x \in \mathcal{N} | f^\infty(x, 0) \leq 0 \text{ a.s.}\} = \{0\},
\]

so the linearity condition is always satisfied.
Example 11 (Shadow price of information) When

\[ f(x, u, \omega) = h(x + u, \omega), \]

the linearity condition means that

\[ \{ x \in \mathcal{N} | h^\infty(x) \leq 0 \text{ -a.s.} \} \]

is linear.
Closedness criteria

Example 12 (Optimal investment) When

\[ f(x, u, \omega) = \begin{cases} v \left( u - \sum_{t=0}^{T-1} x_t \cdot \Delta S_{t+1}(\omega) \right) & \text{if } x_t \in D_t(\omega), \\ +\infty & \text{otherwise} \end{cases} \]

we get

\[ f^\infty(x, u, \omega) = \begin{cases} v^\infty \left( u - \sum_{t=0}^{T-1} x_t \cdot \Delta S_{t+1}(\omega) \right) & \text{if } x_t \in D^\infty_t(\omega) \\ +\infty & \text{otherwise} \end{cases} \]

If \( v \) is nonconstant and \( D_t(\omega) = \mathbb{R}^J \), the linearity condition becomes the no-arbitrage condition

\[ x \in \mathcal{N} : \sum x_t \cdot \Delta S_{t+1} \geq 0 \implies \sum x_t \cdot \Delta S_{t+1} = 0. \]

Example 13 With transaction costs, we get the robust no-arbitrage condition introduced by [Schachermayer, 2004].
Closedness criteria

The linearity condition may hold even under arbitrage.

**Example 14** It holds if $S^\infty_t(x, \omega) > 0$ for $x \notin \mathbb{R}_J$.

**Example 15** In [Çetin and Rogers, 2007] with

$$S_t(x, \omega) = x^0 + s_t(\omega)\psi(x^1)$$

one has $S^\infty_t(x, \omega) = x^0 + s_t(\omega)\psi^\infty(x^1)$. When $\inf \psi' = 0$ and $\sup \psi' = \infty$ we have $\psi^\infty = \delta_{\mathbb{R}_-}$, so the condition in Example 14 holds.

**Example 16** If $S_t(\cdot, \omega) = s_t(\omega) \cdot x$ for a componentwise strictly positive price process $s$ and $D^\infty_t(\omega) \subseteq \mathbb{R}_J^+$ (infinite short selling is prohibited) then linearity condition holds.
**Theorem 6** Assume that $\partial \varphi(u) \neq \emptyset$ and that for every $y \in \partial \varphi(u)$ there exists $v \in \mathcal{N}^\perp$ such that $\varphi^*(y) = Ef^*(v, y)$. Then an $x \in \mathcal{N}$ solves (P) if and only if it is feasible and there exist $y \in \mathcal{Y}$ and $v \in \mathcal{N}^\perp$ such that

$$(v, y) \in \partial f(x, u)$$

$P$-almost surely, or equivalently, if

$$v \in \partial_x l(x, y) \quad \text{and} \quad u \in \partial_y [-l](x, y)$$

$P$-almost surely.
Optimality conditions

**Example 17** In the model with constraints, the optimality condition of Theorem 6 means that

\[ f_j(x) + u_j \leq 0, \]

\[ x \in \arg\min_{z \in \mathbb{R}^n} \left\{ f_0(z) + \sum_{j=1}^{m} y_j f_j(z) - z \cdot v \right\}, \]

\[ y_j f_j(z) = 0 \quad j = 1, \ldots, m, \]

\[ y_j \geq 0 \]

\( P\)-almost surely. [Rockafellar and Wets, 1978] give sufficient conditions for the existence of an optimal \( x \in \mathcal{N}^\infty \) and the corresponding dual variables \( y \in \mathcal{Y} \) and \( v \in \mathcal{N}^\perp \).
Example 18  
Consider the optimal investment problem in liquid markets and assume that $EV$ is finite on $L^p$ and $EV^*$ is proper in $L^q$. Then an $x \in \mathcal{N}$ solves (P) if and only if it is feasible and there exists a martingale density $y$ for $s$ such that

$$y \in \partial V(u - \sum_{t=0}^{T-1} x_t \cdot \Delta s_{t+1}) \quad P\text{-a.s.}$$

much like e.g. in [Schachermayer, 2001] or [Biagini and Frittelli, 2008] in continuous time.
Continuous-time models

- Continuous time models of financial mathematics are often expressed in terms of the stochastic integral (wealth process) of the portfolio process wrt the price process.
- However, the definition of \((x, u) \mapsto Ef(x, u)\) relies on scenariowise description of \(f(x, u)\).
- Our approach is first to consider strategies of bounded variation and then take the lsc hull of the objective \(Ef\)
  - For BV strategies, stochastic integrals are given pathwise and the dual can often be written down explicitly.
  - Taking the lsc hull of \(Ef\) gives rise to stochastic integrals and it does not affect the dual problem.
  - Does require conjugation of integral functionals on spaces of stochastic processes.