OPTIMIZATION MODELING WITH CONVEXITY AND DUALITY

Terry Rockafellar
University of Washington, Seattle
University of Florida, Gainesville

Humboldt University, Berlin —— January, 2009

LECTURE 1

Basic Framework of Optimization

problems of "continuous" rather than "discrete" type

$\mathcal{X}$ some linear space, e.g., $\boldsymbol{R}^n$ or $\mathcal{L}^p$(probability space)

$f : \mathcal{X} \to \bar{\boldsymbol{R}} = [-\infty, \infty]$ some function

$$\operatorname{dom} f = \{x \in \mathcal{X} \,|\, f(x) < \infty\} \quad \text{effective domain}$$
$$\operatorname{epi} f = \{(x, \alpha) \in \mathcal{X} \times \boldsymbol{R} \,|\, f(x) \le \alpha\} \quad \text{epigraph}$$

Abstract model in optimization

$(\mathcal{P})$ minimize $f(x)$ over all $x \in \mathcal{X}$

feasible solutions: $x \in \operatorname{dom} f$

optimal solutions: $x \in \operatorname{argmin} f$ $\quad \operatorname{argmin}(\mathcal{P})$

optimal value: $\inf f$ $\quad \inf(\mathcal{P})$

convex case: f convex, meaning that $\operatorname{epi} f$ is a convex set

$$f((1-\tau)x' + \tau x'') \le (1-\tau)f(x') + \tau f(x'') \text{ for } \tau \in (0,1)$$

Parametric Embedding and Sensitivity

$\mathcal{U}$ = some linear space of perturbations u

$F : \mathcal{X} \times \mathcal{U} \to \bar{R}$ some function with $F(x, 0) = f(x)$

Parameterized model in optimization

$(\mathcal{P}(u))$ minimize $F(x, u)$ over all $x \in \mathcal{X}$

$(\mathcal{P}(0)) = (\mathcal{P})$

convex parameterization: $F(x, u)$ convex in u

full convexity: $F(x, u)$ convex jointly in x and u

Optimal value function

$p(u) = \inf(\mathcal{P}(u)) = \inf_x F(x, u)$, with $p(0) = \inf(\mathcal{P})$

full convexity $\Longrightarrow$ p is convex

sensitivity to perturbations: generalized derivatives of p at 0

Example of Nonlinear Programming

problem model:

minimize $c_0(x)$ over $x \in S$ having $c_i(x) \leq 0$ for $i = 1, \ldots, m$

$S \subset \mathcal{X}, \quad c_i : S \to R$ for $i = 0, 1, \ldots, m$

corresponding objective in abstract format:

$f(x) = c_0(x)$ if $x \in S$ and $c_i(x) \leq 0$ for $i = 1, \ldots, m$

but otherwise $f(x) = \infty$

canonical parameterization: $u = (u_1, \ldots, u_m)$

$F(x, u) = c_0(x)$ if $x \in S$ and $c_i(x) + u_i \leq 0$ for $i = 1, \ldots, m$

but otherwise $F(x, u) = \infty$

Observations:

- f is convex if $S =$ convex set and each $c_i =$ convex function
- $F(x, u)$ is always convex in u
- $F(x, u)$ is jointly convex in x and u when f is convex.

Example of Composite Objectives

problem model: minimize $\theta(g_1(x), \ldots, g_d(x))$ over all $x \in S$

$S \subset \mathcal{X}$, $c_i : \mathcal{X} \to \boldsymbol{R}$, $\theta : \boldsymbol{R}^d \to (-\infty, \infty]$ convex nondecreasing

corresponding objective function in abstract format:

$$f(x) = \theta(g_1(x), \ldots, g_d(x)) \text{ if } x \in S$$

but otherwise $f(x) = \infty$

canonical parameterization: $u = (u_1, \ldots, u_d)$

$$F(x, u) = \theta(g_1(x) + u_1, \ldots, g_d(x) + u_d) \text{ if } x \in SX$$

but otherwise $F(x, u) = \infty$

Observations:

- f is convex when $S =$ convex set, each $g_i =$ convex function
- $F(x, u)$ is always convex in u
- $F(x, u)$ is jointly convex in x and u when f is convex.

Example of Stochastic Programming

$(\Omega, \mathcal{F}, P) =$ probability space of future states ω

One-stage model

minimize $\Phi(x_0) = E_\omega\{f(x_0, \omega)\}$ over all $x_0 \in \mathcal{X}_0$

$f : \mathcal{X}_0 \times \Omega \to \bar{R}$ incorporates constraints!

$\Phi(x_0) < \infty$ will require $f(x_0, \omega) < \infty$ a.s. in ω

(various technicalities involving measurability need attention)

Two-stage model

minimize $\Phi(x_0, x_1(\cdot)) = E_\omega\{f(x_0, x_1(\omega), \omega)\}$ over all

$x_0 \in \mathcal{X}_0$ and [measurable] mappings $x_1(\cdot) : \Omega \to \mathcal{X}_1$

$x_1(\omega) =$ recourse decision

The expectation functionals Φ are special **integral functionals**

Φ inherits convexity from the integrand f

Lagrangians and Dual Problems

primal problem $(\mathcal{P})$: minimize $f(x)$ over $x \in \mathcal{X}$

Lagrangian for $(\mathcal{P})$ and a multiplier space $\mathcal{Y}$

any function L on $\mathcal{X} \times \mathcal{Y}$ having

$$f(x) = \sup_{y \in \mathcal{Y}} L(x, y) \text{ for all } x \in \mathcal{X}$$

let $g(y) = \inf_{x \in \mathcal{X}} L(x, y)$ for all $y \in \mathcal{Y}$

dual problem $(\mathcal{D})$: maximize $g(y)$ over all $y \in \mathcal{Y}$,

Basic primal-dual relationships

(a) $\inf(\mathcal{P}) \geq \sup(\mathcal{D})$ always

(b) $\left[\, \inf(\mathcal{P}) = \sup(\mathcal{D}),\ \ \bar{x} \in \operatorname{argmin}(\mathcal{P}),\ \ \bar{y} \in \operatorname{argmax}(\mathcal{D}) \,\right]$

$\Longleftrightarrow \left[\, \inf_x L(x, \bar{y}) = L(\bar{x}, \bar{y}) = \sup_y L(\bar{x}, y) \,\right]$ saddle point

saddle point existence: unlikely unless $L(x, y)$ is convex-concave

Paired Spaces for Developing Duality

linear spaces $\mathcal{U}$ and $\mathcal{Y}$, with bilinear form $\langle u, y\rangle$ on $\mathcal{U} \times \mathcal{Y}$

Compatible topologies

the continuous linear functionals on $\mathcal{U}$ are $u \longrightarrow \langle u, y\rangle$ for $y \in \mathcal{Y}$
the continuous linear functionals on $\mathcal{Y}$ are $y \longrightarrow \langle u, y\rangle$ for $u \in \mathcal{U}$

Examples:

- $\mathcal{U} = \boldsymbol{R}^m, \ \mathcal{Y} = \boldsymbol{R}^m, \ \langle u, y\rangle = u{\cdot}y = \sum_{i=1}^m u_i y_i$ usual topology
- $\mathcal{U} = \mathcal{L}^p_m(\Omega, \mathcal{F}, P), \ \mathcal{Y} = \mathcal{L}^q_m(\Omega, \mathcal{F}, P)$, usual pairing,
  with $\langle u, y\rangle = E\{u{\cdot}y\} = \int_\Omega \sum_{i=1}^m u_i(\omega) y_i(\omega) dP(\omega)$
  the norm topologies, except for $\mathcal{L}^\infty$ the weak* topology
- the weak topologies $\sigma(\mathcal{U}, \mathcal{Y})$ on $\mathcal{U}$ and $\sigma(\mathcal{Y}, \mathcal{U})$ on $\mathcal{Y}$

Note: the **closed convex** sets and **lsc convex** functions (lower semicontinuous) are **the same in all compatible topologies**

Conjugate Convex Functions

$\mathcal{U}$ and $\mathcal{Y}$: paired linear spaces with compatible topologies

Legendre-Fenchel transform

$\varphi : \mathcal{U} \to \bar{R}$ any function
$\varphi^* : \mathcal{Y} \to \bar{R}$ its conjugate, $\varphi^*(y) = \sup_u \{\langle u, y\rangle - \varphi(u)\}$
$\varphi^{**} : \mathcal{U} \to \bar{R}$ its biconjugate, $\varphi^{**}(u) = \sup_y \{\langle u, y\rangle - \varphi^*(y)\}$

Closed* convex functions (lsc and $> -\infty$, unless $\equiv -\infty$)

- φ^* is a closed* convex function
- φ^{**} is the largest closed* convex function $\leq \varphi$

Conjugacy correspondence

The **closed* convex** functions φ on $\mathcal{U}$ and ψ on $\mathcal{Y}$ correspond **one-to-one** to each other under: $\psi = \varphi^*, \quad \varphi = \psi^*$

The constant functions ∞ and $-\infty$ are conjugate to each other

Conjugate Duality Scheme in Optimization

$\mathcal{U}$ and $\mathcal{Y}$: paired linear spaces with compatible topologies

For the problem $(\mathcal{P})$ of minimizing $f(x)$ over $x \in \mathcal{X}$, consider

- parameterizations $F : \mathcal{X} \times \mathcal{U} \to \bar{R}$ with $F(x, \cdot)$ closed* convex
- Lagrangians $L : \mathcal{X} \times \mathcal{Y} \to \bar{R}$ with $-L(x, \cdot)$ closed* convex

Parameterizations versus Lagrangians

Such F and L correspond to each other **one-to-one** under

$$L(x, y) = \inf_u \{F(x, u) - \langle u, y\rangle\}, \quad F(x, u) = \sup_u \{L(x, y) + \langle u, y\rangle\}$$

$F(x, u)$ convex in (x, u) $\iff$ $L(x, y)$ concave in y

Nonlinear programming example: $u \in R^m$, $y \in R^m$

$$F(x, u) = c_0(x) \text{ if } x \in S \text{ and } c_i(x) + u_i \leq 0 \text{ for } i = 1, \ldots, m$$

but otherwise $F(x, u) = \infty$

$$L(x, y) = c_0(x) + y_1 c_1(x) + \cdots + y_m c_m(x) \text{ if } x \in S,\ y \geq 0$$

and $= \infty$ if $x \notin S$, $y \geq 0$, but $= -\infty$ if $y \not\geq 0$

Main Results for the Conjugate Duality Scheme

$\mathcal{U}$ and $\mathcal{Y}$: paired linear spaces with compatible topologies

Lagrangian $L(x, y) \quad \leftrightarrow \quad$ parameterization $F(x, u)$

$(\mathcal{P})$ minimize $f(x)$ over $x \in X$ where $f(x) = \sup_y L(x, y)$

$(\mathcal{D})$ maximize $g(y)$ over $y \in Y$ where $g(y) = \inf_x L(x, y)$

Optimal value function:

$$p(u) = \inf_x F(x, u) = \inf(\mathcal{P}(u)) \qquad \text{where } F(x, 0) = f(x)$$

Characterization of primal-dual optimal values and solutions

(a) $\inf(\mathcal{P}) = p(0), \quad \sup(\mathcal{D}) = p^{**}(0)$

(b) $(\bar{x}, \bar{y})$ is a saddle point of $L(x, y)$ if and only if

$\bar{x} \in \mathrm{argmin}(\mathcal{P})$ and $p(u) \geq p(0) + \langle u, \bar{y} \rangle$ for all $u \in \mathcal{U}$

Key question: when does there exist $\bar{y}$ with this relation to p at 0?

Subgradients and Directional Derivatives

Subgradients of convex analysis

For $\varphi : \mathcal{U} \to \bar{R}$, $\varphi \not\equiv \infty$, $u \in \mathcal{U}$, $y \in \mathcal{Y}$:

$$y \in \partial\varphi(u) \quad \text{means} \quad \varphi(u+w) \geq \varphi(u) + \langle w, y\rangle \text{ for all } w \in \mathcal{U}$$

Directional derivatives of convex functions

For φ convex on $\mathcal{U}$, finite at $\bar{u}$, bounded above around $\bar{u}$:

(a) $\varphi'(\bar{u}; w) = \lim_{\tau \to 0^+} \dfrac{\varphi(\bar{u} + \tau w) - \varphi(\bar{u})}{\tau}$ is finite, convex in w

(b) $\varphi'(\bar{u}; w) = \max\left\{ \langle w, y\rangle \,\middle|\, y \in \partial\varphi(\bar{u})\right\}$

(c) for R^n: $\partial\varphi(\bar{u}) = \{\bar{y}\} \iff \varphi$ diff. at $\bar{u}$ with $\bar{y} = \nabla\varphi(\bar{u})$

Relation to conjugacy

For conjugate functions φ on $\mathcal{U}$ and ψ on $\mathcal{V}$, not $\equiv \infty$ or $\equiv -\infty$:

(a) $\varphi(u) + \varphi(y) \geq \langle u, y\rangle$ for all $u \in \mathcal{U}$ and $y \in \mathcal{Y}$

(b) equality holds for $u, y \iff y \in \partial\varphi(u) \iff u \in \partial\psi(y)$

Fenchel-Type Duality Schemes

$\mathcal{U} \leftrightarrow \mathcal{Y}$, $\mathcal{X} \leftrightarrow \mathcal{V}$: paired linear spaces with compatible topologies

proper lsc convex h on $\mathcal{X}$, k on $\mathcal{U}$, conjugates h^* on $\mathcal{V}$, k^* on $\mathcal{Y}$

$c \in \mathcal{V}$, $b \in \mathcal{U}$, continuous linear $A : \mathcal{X} \to \mathcal{U}$, adjoint $A^* : \mathcal{Y} \to \mathcal{V}$

primal $(\mathcal{P})$ $\min f(x) = \langle c, x \rangle + h(x) + k(b - Ax)$ over $x \in \mathcal{X}$

dual $(\mathcal{D})$ $\max g(y) = \langle b, y \rangle - k^*(y) - h^*(A^*y - c)$ over $y \in \mathcal{Y}$

Lagrangian: $L(x, y) = \langle c, x \rangle + h(x) + \langle b, y \rangle - k^*(y) - \langle Ax, y \rangle$

feasibility in $(\mathcal{P})$ $\iff$ $b \in [A\,\mathrm{dom}\,h + \mathrm{dom}\,k]$

feasibllity in $(\mathcal{D})$ $\iff$ $c \in [A^*\,\mathrm{dom}\,k^* - \mathrm{dom}\,h^*]$

Duality Theorem

Suppose $\mathcal{U}$ and $\mathcal{V}$ are Banach (in the compatible topologies!)

(a) $\inf(\mathcal{P}) = \max(\mathcal{D}) < \infty$ if $b \in \mathrm{int}[A\,\mathrm{dom}\,h + \mathrm{dom}\,k]$

(b) $\min(\mathcal{P}) = \sup(\mathcal{D}) > -\infty$ if $c \in \mathrm{int}[A^*\,\mathrm{dom}\,k^* - \mathrm{dom}\,h^*]$

Some General References

[1] R. T. Rockafellar (1974), *Conjugate Duality and Optimization*, No. 16 in the Conference Board of Math. Sciences Series, SIAM Publications, Philadelphia (74 pages)

[2] R. T. Rockafellar (1970), *Convex Analysis*, Princeton University Press, Princeton, New Jersey (available from 1997 also in paperback in the series Princeton Landmarks in Mathematics).

[3] R. T. Rockafellar, R. J.-B. Wets (1998, 2005), *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften 317, Springer-Verlag, Berlin (second printing, with corrections: 2005)

[4] R. T. Rockafellar (1999), "Extended nonlinear programming," in *Nonlinear Optimization and Related Topics* (G. Di Pillo and F. Giannessi, eds,), Kluwer, 381-399 downloadable

[5] R. T. Rockafellar (1993), "Lagrange multipliers and optimality," *SIAM Review* 35, 183–238

DOWNLOADS

website: www.math.washington.edu/$\sim$rtr/mypage.html

Available besides [4] and some other relatively recent papers:

- Course lecture notes on *Fundamentals of Optimization*

Very introductory material in finite dimensions, which nonetheless covers geometric nonsmooth analysis and optimality conditions in terms of normal cones, as well as properties of polyhedrality

- Course lecture notes on *Optimization Under Uncertainty*,

The basics of traditional stochastic programming, without use of "risk measures," but with duality and a build-up to multistage models in a framework of scenarious and decomposition