Convex Optimization

Kontonis Vasilis

3 March, 2017

Corelab, NTUA

Contents

Vector Optimization

Dual Inequalities

Dual Cone
Let X be a vector space and X^* be its dual

- If $K \subseteq X$ is a cone then its dual cone is the set

$$K^* = \{y \in X^* \mid y^T x \geqslant 0, \text{ for all } x \in K\}$$

- $(\mathbb{R}^n_+)^* = \mathbb{R}^n_+$
- $(S^n_+)^* = S^n_+$
- K^* is always convex.
- K proper $\implies$ K^* proper.

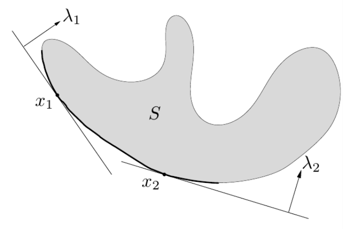
Dual Inequalities
$x \preccurlyeq_K y \Leftrightarrow \lambda^T x \leqslant \lambda^T y$ for all $\lambda \succcurlyeq_{K^*} 0$.

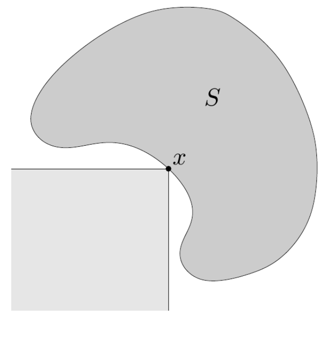
Minimum Element
x is minimum in $S \Leftrightarrow$ for all $\lambda \succ_{K^*} 0$, x is the unique minimizer of $\lambda^T z$ over $z \in S \Leftrightarrow$ The hyperplane $\{z \mid \lambda^T(z - x) = 0\}$ is a strict supporting hyperplane to S at x for all $\lambda \in K^*$.

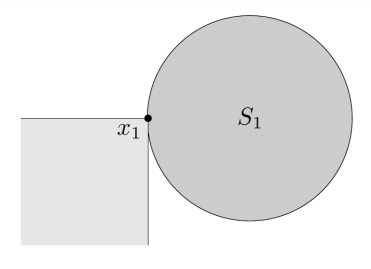
Minimal Elements

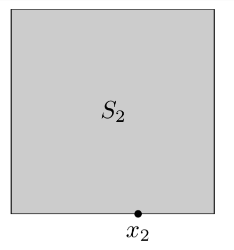
- If $\lambda^T \succ_{K^*} 0$ and x minimizes $\lambda^T z$ over $z \in S$, then x is minimal.
- If S is convex, for any minimal element x there exists nonzero $\lambda \succcurlyeq_{K^*} 0$ s.t. x minimizes $\lambda^T z$ over $z \in S$.

Counterexamples









Convex Vector Optimization Problem

Let $f_0 : \mathbb{R}^n \to \mathbb{R}^q$, $K \subseteq \mathbb{R}^q$ a proper cone.

$$\begin{array}{ll} \text{minimize (with respect to } K) & f_0(x) \\ \text{subject to} & f_i(x) \leqslant 0 \\ & h_i(x) = 0 \end{array}$$

- f_0 is K-convex.
- f_i are convex.
- h_i are affine.

A point x^* is optimal iff it is feasible and

$$f_0(D) \subseteq f_0(x^*) + K$$

Scalarization

Pareto Optimal Points

- A point x is Pareto optimal iff it is feasible and $(f_0(x) - K) \cap f_0(D) = \{f_0(x)\}$
- The set of Pareto optimal values, $\mathcal{P}$ satisfies $\mathcal{P} \subseteq f_0(D) \cap \partial f_0(D)$

Scalarization

Let $\lambda \succeq_{K^*} 0$ be the weight vector.

$$\begin{aligned} \text{minimize} \quad & \lambda^T f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0 \\ & h_i(x) = 0 \end{aligned}$$

If the problem is convex then **every** pareto optimal point is attainable via scalarization.

Minimal Matrix Upper Bound

$$\begin{aligned} \text{minimize (w.r.t } S_+^n) \quad & X \\ \text{subject to} \quad & X \succcurlyeq A_i, \ i = 1, \dots, m \end{aligned}$$

Let $W \in S_{++}^n$ and form the equivalent **SDP**

$$\begin{aligned} \text{minimize (w.r.t } S_+^n) \quad & \mathrm{tr}(WX) \\ \text{subject to} \quad & X \succcurlyeq A_i, \ i = 1, \dots, m \end{aligned}$$

Ellipsoids and Positive Definiteness
$\mathcal{E}_A = \{u \mid u^T A^{-1} u \leqslant 1\}$
$A \preccurlyeq B \Leftrightarrow \mathcal{E}_A \subseteq \mathcal{E}_B$

Duality

Langrangian

Langrangian
$L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, with $\mathrm{dom} L = D \times \mathbb{R}^m \times \mathbb{R}^p$.

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x)$$

Dual function
$g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

$$g(\lambda, \mu) = \inf_{x \in D} L(x, \lambda, \mu)$$

Dual function for $\lambda \geqslant 0$ underestimates the optimal value $g(\lambda, \mu) \leqslant p^*$.

Multicriterion Interpretation

Primal Problem without equality constraints:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leqslant 0, \ i = 1, \dots, m \end{aligned}$$

Scalarization of the multicreterion problem:

$$\text{minimize} \quad F(x) = (f_0(x), \ f_1(x), \dots, f_m(x))$$

Take $\widetilde{\lambda} = (1, \lambda)$ and then minimize

$$\widetilde{\lambda}^T F(x) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$$

which is the Langrangian of the Primal Problem.

Nonconvex QCQP

Let $A \in S^n, A \not\geq 0, b \in \mathbb{R}^n$.

$$\begin{array}{ll} \text{maximize} & x^T A x + 2b^T x \\ \text{subject to} & x^T x \leqslant 1 \end{array}$$

Langrangian:

$$L(x, \lambda) = x^T A x + 2b^T x + \lambda(x^T x - 1) = x^T(A + \lambda I)x + 2b^T x - \lambda$$

Dual Function:

$$g(\lambda) = \begin{cases} -b^T(A + \lambda I)^{\dagger} b - \lambda, & A + \lambda I \geqslant 0, \quad b \in \mathcal{R}(A + \lambda I) \\ -\infty & \text{otherwise} \end{cases}$$

Nonconvex QCQP

Dual Problem

$$\begin{aligned} &\text{maximize} \quad -b^T(A+\lambda I)^{\dagger}b-\lambda \\ &\text{subject to} \;\; A+\lambda I \succeq 0, b \in \mathcal{R}(A+\lambda I) \end{aligned}$$

We can find an equivalent **concave** problem

$$\begin{aligned} &\text{maximize} \quad -\sum_{i=1}^{n}\frac{(q_i^Tb)^2}{\lambda_i+\lambda}-\lambda \\ &\text{subject to} \;\; \lambda \geq -\lambda_{\min}(A) \end{aligned}$$

For these problems strong duality obtains.

Rayleigh Quotient

Let $A \in S^n$

$$\text{maximize} \quad \frac{x^T A x}{x^T x}$$

Equivalent problem:

$$\begin{aligned} \text{maximize} \quad & x^T A x \\ \text{subject to} \quad & x^T x \leqslant 1 \end{aligned}$$

Lagrangian: $L(x, \mu) = x^T A x + \lambda(x^T x - 1)$

Derivative

Let E, F be **Banach Spaces**, that is complete normed spaces.

Derivative is a Linear Map
Let U be open in E, and let $x \in U$. Let $f : U \to F$ be a map. f is **differentiable** at x if there exists a continuous linear map $\lambda : E \to F$ and a map ψ defined for all sufficiently small h in E, with values in F, such that

$$\lim_{h \to 0} \psi(h) = 0, \text{ and } f(x+h) = f(x) + \lambda(h) + |h|\psi(h).$$

log(det(X))

$$f(X) : S^n_{++} \to \mathbb{R}, \;\; f(X) = \log\det(X)$$

$$\begin{aligned}
\log\det(X+H) &= \log\det(X+H) \\
&= \log\det\left(X^{1/2}(I + X^{-1/2}HX^{-1/2})X^{1/2}\right) \\
&= \log\det X + \log\det(I + X^{-1/2}HX^{-1/2}) \\
&= \log\det X + \sum_{i=1}^{n} \log(1+\lambda_i) \\
&\simeq \log\det X + \sum_{i=1}^{n} \lambda_i \\
&= \log\det X + \mathrm{tr}(X^{-1/2}HX^{-1/2}) \\
&= \log\det X + \mathrm{tr}(X^{-1}H)
\end{aligned}$$

$$\nabla f(X) = X^{-1}$$

Conjugate of logdet

Conjugate function:

$$f^*(y) = \sup_{x \in D} (y^T x - f(x))$$

$f(X) = \log\det X^{-1}, X \in S^n_{++}$
The conjugate of f is

$$f^*(Y) = \sup_{X > 0} (\mathrm{tr}(YX) + \log\det X)$$

- $\mathrm{tr}(YX) + \log\det X$ is unbounded if $Y \nless 0$.
- If $Y < 0$ then setting the gradient with respect to X to zero yields $X_0 = -Y^{-1}$

$f^*(Y) = \log\det(-Y)^{-1} - n = -\log\det(-Y) - n$
$\mathrm{dom}\ f^* = -S^n_{++}$

Dual of Affine Constraints

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \leqslant b \\ & Cx = d \end{array}$$

$$\begin{aligned} g(\lambda, \mu) &= \inf_x \left(f_0(x) + \lambda^T(Ax - b) + \mu^T(Cx - d)\right) \\ &= -b^T\lambda - d^T\mu + \inf_x \left(f_0(x) + (A^T\lambda - C^T\mu)\right) \\ &= -b^T\lambda - d^T\mu - f_0^*(-A^T\lambda - C^T\mu) \end{aligned}$$

with $\operatorname{dom} g = \{(\lambda, \mu) \mid -A^T\lambda - C^T\mu \in \operatorname{dom} f_0^*\}$

Minimum Volume Covering Ellipsoid

Primal

$$\begin{aligned} &\text{minimize} \quad f_0(X) = \log\det(X^{-1}) \\ &\text{subject to} \;\; a_i^T X a_i \leqslant 1, \; i = 1, \dots, m \end{aligned}$$

$a_i^T X a_i \Leftrightarrow \text{tr}(a_i a_i^T X) \leqslant 1$

Dual Function

$$g(\lambda, \nu) = \begin{cases} \log\det\left(\sum_{i=1}^m \lambda_i a_i a_i^T\right) - \mathbf{1}^T\lambda + n, & \sum_{i=1}^m \lambda_i a_i a_i^T > 0 \\ -\infty, & \text{otherwise} \end{cases}$$

Dual

$$\begin{aligned} &\text{minimize} \quad \log\det\left(\sum_{i=1}^m \lambda_i a_i a_i^T\right) - \mathbf{1}^T\lambda + n \\ &\text{subject to} \;\; \lambda \geqslant 0 \end{aligned}$$

The weaker Slater condition is satisfied ($\exists X \in S_{++}^n, a_i^T X a_i \leqslant 1, i \in [m]$) and therefore Strong Duality obtains.

The Perturbed Problem

The perturbed version of the convex problem:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leqslant u_i, \; i = 1, \ldots, m \\ & h_i(x) = v_i, \; i = 1, \ldots, p \end{aligned}$$

The optimal value:

$$p^*(u, v) = \inf\{f_0(x) \mid \exists x \in D, f_i(x) \leqslant u_i, h_i(x) = v_i\}$$

- The optimal value of the unperturbed problem is $p^*(0, 0) = p^*$
- When the perturbations result in infeasibility we have $p^*(u, v) = \infty$.
- $p^*(u, v)$ is convex when the original problem is convex.

A Global Inequality

Assume that the original problem is **convex** and Slater's condition is satisfied.
Let (λ^*, μ^*) be optimal for the dual of the original problem. Then

$$p^*(u, v) \geqslant p^*(0, 0) - {\lambda^*}^T u - {\mu^*}^T v$$

Proof.

$$\begin{aligned} p^*(0, 0) &= g(\lambda^*, \mu^*) \\ &\leqslant f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \mu_i^* h_i(x) \\ &\leqslant f_0(x) + {\lambda^*}^T u + {\mu^*}^T v \end{aligned}$$

□

Interpretation of the Global Inequality

$p^*(u, v) \geqslant p^*(0, 0) - {\lambda^*}^T u - {\mu^*}^T v$

- λ_i^* is large, $u_i < 0$ then $p^*(u, v)$ will increase greatly.
- μ_i^* is large and positive, $v_i < 0$ OR μ_i^* is large and negative, $v_i > 0$ then $p^*(u, v)$ will increase greatly.
- If λ_i^* is small, $u_i > 0$ then $p^*(u, v)$ will not decrease too much.
- If μ_i^* is small and positive ,$v_i > 0$ OR μ_i^* is small and negative and $v_i < 0$
  then $p^*(u, v)$ will not decrease too much.

These results are **not symmetric** with respect to tightening or loosening a constraint.

Local Sensitivity Analysis

Assume strong duality and differentiability of $p^*(u, v)$ at $(0, 0)$.

$$\lambda_i^* = -\frac{\partial p^*}{\partial u_i}\bigg|_{(0,0)}, \quad \mu_i^* = -\frac{\partial p^*}{\partial v_i}\bigg|_{(0,0)}$$

Differentiability of p^* allows a symmetric sensitivity result.

Proof.

$$\frac{\partial p^*}{\partial u_i}\bigg|_{(0,0)} = \lim_{t\to 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t}$$

From the global inequality we have

$$\frac{p(u,v) - p^*(0,0)}{t} \geqslant -\lambda_i \text{ if } t > 0 \text{ and } \frac{p(u,v) - p^*(0,0)}{t} \leqslant -\lambda_i \text{ if } t < 0$$

□

Duality in SDP

Primal SDP:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & x_1 F_1 + \ldots + x_n F_n + G \preccurlyeq 0 \end{aligned}$$

Then

$$\begin{aligned} L(x, Z) &= c^T x + \operatorname{tr}((x_1 F_1 + \ldots + x_n F_n + G) Z) \\ &= x_1 (c_1 + \operatorname{tr}(F_1 Z)) + \ldots + x_n (c_n + \operatorname{tr}(F_n Z)) + \operatorname{tr}(GZ) \end{aligned}$$

Dual function:

$$g(Z) = \inf_x L(x, Z) = \begin{cases} \operatorname{tr}(GZ), & \operatorname{tr}(F_i Z) + c_i = 0,\ i = 1, \ldots, n \\ -\infty, & \text{otherwise} \end{cases}$$

Duality in SDP

Dual Problem:

$$\begin{aligned} \text{minimize} \quad & \operatorname{tr}(GZ) \\ \text{subject to} \quad & \operatorname{tr}(F_i Z) + c_i = 0,\ i = 1, \ldots, n \\ & Z \geqslant 0 \end{aligned}$$

Strong Duality obtains if the SDP is strictly feasible, namely there exists an x with

$$x_1 F_1 + \ldots + x_n F_n + G < 0$$

Questions?

References i

- S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan.
  Linear Matrix Inequalities in System and Control Theory.
  Society for Industrial and Applied Mathematics, Philadelphia, June 1997.
- S. Boyd and L. Vandenberghe.
  Convex Optimization.
  Cambridge University Press, Cambridge, UK ; New York, Mar. 2004.