EE/ACM 150 - Applications of Convex Optimization in Signal Processing and Communications
Lecture 11

Andre Tkacenko

Signal Processing Research Group
Jet Propulsion Laboratory

May 8, 2012

Outline

1. Lagrangian Duality
   - The Lagrangian and Lagrange Dual Function
     - Examples
   - Lagrange Dual Problem
   - Weak & Strong Duality
     - Examples
   - Geometric Interpretation

The Lagrangian

Recall the standard form optimization problem (which need not be convex).

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, m \quad, \\ & h_i(\mathbf{x}) = 0\,,\ i = 1, \ldots, p \end{array}$$

with variable $\mathbf{x} \in \mathbb{R}^n$, domain $\mathcal{D} = \bigcap_{i=0}^{m} \operatorname{dom}(f_i) \cap \bigcap_{i=1}^{p} \operatorname{dom}(h_i)$, and optimal value $p^\star$.
The *Lagrangian* is a function that combines the objective with the constraint functions.

Lagrangian

The Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, with $\operatorname{dom}(L) = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$, is defined as

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \triangleq f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \ .$$

- It is a weighted sum of the objective and constraint functions.
- The quantity λ_i is the *Lagrange multiplier* associated with $f_i(\mathbf{x}) \leq 0$.
- The quantity ν_i is the *Lagrange multiplier* associated with $h_i(\mathbf{x}) = 0$.
- The vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ are called the *dual variables* associated with the problem.

The Lagrange Dual Function

The minimum value of the Lagrangian over $\mathbf{x}$ is known as the *Lagrange dual function*.

Lagrange Dual Function

The Lagrange dual function $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is defined as

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \triangleq \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x} \in \mathcal{D}} \left\{ f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \right\} .$$

- As the dual function g is the pointwise infimum of a family of affine functions of $(\boldsymbol{\lambda}, \boldsymbol{\nu})$, it is always *concave*, even if the original or primal problem is not convex.
- When the Lagrangian L is unbounded below in $\mathbf{x}$, the dual function g takes on the value $-\infty$.

Lower Bound Property: If $\boldsymbol{\lambda} \succeq \mathbf{0}$, then $g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq p^\star$.

Proof: If $\widetilde{\mathbf{x}}$ is feasible and $\boldsymbol{\lambda} \succeq \mathbf{0}$, then we have

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \leq L(\widetilde{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\nu}) \leq f_0(\widetilde{\mathbf{x}}) .$$

Minimizing over all feasible $\widetilde{\mathbf{x}}$ gives $g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq f_0(\mathbf{x}^\star) = p^\star$.

Least-Norm Solution of Linear Equations

$$\begin{array}{ll} \text{minimize} & \mathbf{x}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array} .$$

Dual function:

- The Lagrangian is $L(\mathbf{x}, \boldsymbol{\nu}) = \mathbf{x}^T\mathbf{x} + \boldsymbol{\nu}^T (\mathbf{A}\mathbf{x} - \mathbf{b})$.
- To minimize L over $\mathbf{x}$, simply set the gradient with respect to $\mathbf{x}$ equal to zero. This yields

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\nu}) = 2\mathbf{x} + \mathbf{A}^T\boldsymbol{\nu} = \mathbf{0} \Longrightarrow \mathbf{x} = -(1/2)\,\mathbf{A}^T\boldsymbol{\nu} \,.$$

- Plug in this value of $\mathbf{x}$ into L to obtain g. This results in

$$g(\boldsymbol{\nu}) = L\big(-(1/2)\,\mathbf{A}^T\boldsymbol{\nu}, \boldsymbol{\nu}\big) = -(1/4)\,\boldsymbol{\nu}^T\mathbf{A}\mathbf{A}^T\boldsymbol{\nu} - \mathbf{b}^T\boldsymbol{\nu} \,,$$

  which is a concave function of $\boldsymbol{\nu}$.

Lower bound property: $p^\star \geq -(1/4)\,\boldsymbol{\nu}^T\mathbf{A}\mathbf{A}^T\boldsymbol{\nu} - \mathbf{b}^T\boldsymbol{\nu}$ for all $\boldsymbol{\nu}$.

Standard Form LP

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b}\,,\ \mathbf{x} \succeq \mathbf{0} \end{array}.$$

Dual function:

- The Lagrangian is

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) &= \mathbf{c}^T\mathbf{x} + \boldsymbol{\nu}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) - \boldsymbol{\lambda}^T\mathbf{x}\,, \\ &= -\mathbf{b}^T\boldsymbol{\nu} + \left(\mathbf{c} + \mathbf{A}^T\boldsymbol{\nu} - \boldsymbol{\lambda}\right)^T\mathbf{x}\,. \end{aligned}$$

- Note that L is affine in $\mathbf{x}$. Hence, we have

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = \begin{cases} -\mathbf{b}^T\boldsymbol{\nu}\,, & \mathbf{A}^T\boldsymbol{\nu} - \boldsymbol{\lambda} + \mathbf{c} = \mathbf{0} \\ -\infty\,, & \text{otherwise} \end{cases}.$$

  Here, g is linear on an affine domain $\left\{(\boldsymbol{\lambda}, \boldsymbol{\nu}) : \mathbf{A}^T\boldsymbol{\nu} - \boldsymbol{\lambda} + \mathbf{c} = \mathbf{0}\right\}$, and thus is concave.

Lower bound property: $p^\star \geq -\mathbf{b}^T\boldsymbol{\nu}$ if $\mathbf{A}^T\boldsymbol{\nu} + \mathbf{c} \succeq \mathbf{0}$.

Equality Constrained Norm Minimization

$$\begin{array}{ll} \text{minimize} & ||\mathbf{x}|| \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array}.$$

Dual function:

- The Lagrangian is $L(\mathbf{x}, \boldsymbol{\nu}) = ||\mathbf{x}|| - \boldsymbol{\nu}^T (\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{b}^T \boldsymbol{\nu} + ||\mathbf{x}|| - \left(\mathbf{A}^T \boldsymbol{\nu}\right)^T \mathbf{x}$.
- The dual function is

$$g(\boldsymbol{\nu}) = \mathbf{b}^T \boldsymbol{\nu} + \inf_{\mathbf{x}} \left\{ ||\mathbf{x}|| - \left(\mathbf{A}^T \boldsymbol{\nu}\right)^T \mathbf{x} \right\} = \begin{cases} \mathbf{b}^T \boldsymbol{\nu}, & \left|\left|\mathbf{A}^T \boldsymbol{\nu}\right|\right|_{\circledast} \leq 1 \\ -\infty, & \text{otherwise} \end{cases},$$

where recall that $||\mathbf{v}||_{\circledast} = \sup_{||\mathbf{u}|| \leq 1} \left\{ \mathbf{u}^T \mathbf{v} \right\}$ is the *dual norm* of $||\cdot||$.

Proof: This follows from the fact that

$$\inf_{\mathbf{x}} \left\{ ||\mathbf{x}|| - \mathbf{y}^T \mathbf{x} \right\} = \begin{cases} 0, & ||\mathbf{y}||_{\circledast} \leq 1 \\ -\infty, & ||\mathbf{y}||_{\circledast} > 1 \end{cases}.$$

  - If $||\mathbf{y}||_{\circledast} \leq 1$, then we have $||\mathbf{x}|| \geq ||\mathbf{x}|| \cdot ||\mathbf{y}||_{\circledast} \geq \mathbf{y}^T \mathbf{x}$ for all $\mathbf{x}$. Hence, $||\mathbf{x}|| - \mathbf{y}^T \mathbf{x} \geq 0$ for all $\mathbf{x}$, with equality if $\mathbf{x} = \mathbf{0}$.
  - If $||\mathbf{y}||_{\circledast} > 1$, then choose $\mathbf{x} = t\mathbf{u}$, where $t > 0$ and $\mathbf{u}$ is such that $||\mathbf{u}|| \leq 1$ and $\mathbf{u}^T \mathbf{y} = ||\mathbf{y}||_{\circledast} > 1$. Note that such a $\mathbf{u}$ always exists. In this case, we get

$$||\mathbf{x}|| - \mathbf{y}^T \mathbf{x} = t \left(||\mathbf{u}|| - ||\mathbf{y}||_{\circledast} \right) \to -\infty \text{ as } t \to \infty.$$

Lower bound property: $p^\star \geq \mathbf{b}^T \boldsymbol{\nu}$ if $\left|\left|\mathbf{A}^T \boldsymbol{\nu}\right|\right|_{\circledast} \leq 1$.

Two-Way Partitioning Problem

$$\begin{array}{ll} \text{minimize} & \mathbf{x}^T \mathbf{W} \mathbf{x} \\ \text{subject to} & x_i^2 = 1\,,\ i = 1, \ldots, n \end{array} .$$

Here, $\mathbf{W} \in \mathbb{S}^n$.

- This is a nonconvex, NP-complete problem; the feasible set contains 2^n discrete points.
- The interpretation of the problem is to partition $\{1, \ldots, n\}$ into two sets, where $W_{k,\ell}$ is the cost of assigning k and ℓ to the same set, while $-W_{k,\ell}$ is the cost of assigning k and ℓ to different sets.

Dual function:

$$\begin{aligned} g(\boldsymbol{\nu}) &= \inf_{\mathbf{x}} \left\{ \mathbf{x}^T \mathbf{W} \mathbf{x} + \sum_{i=1}^{n} \nu_i \left(x_i^2 - 1 \right) \right\} = \inf_{\mathbf{x}} \left\{ \mathbf{x}^T \left(\mathbf{W} + \operatorname{diag}(\boldsymbol{\nu}) \right) \mathbf{x} - \mathbf{1}^T \boldsymbol{\nu} \right\} , \\ &= \begin{cases} -\mathbf{1}^T \boldsymbol{\nu}\,, & \mathbf{W} + \operatorname{diag}(\boldsymbol{\nu}) \succeq \mathbf{0} \\ -\infty\,, & \text{otherwise} \end{cases} . \end{aligned}$$

Lower bound property: $p^\star \geq -\mathbf{1}^T \boldsymbol{\nu}$ if $\mathbf{W} + \operatorname{diag}(\boldsymbol{\nu}) \succeq \mathbf{0}$.

Example: $\boldsymbol{\nu} = -\lambda_{\min}(\mathbf{W})\, \mathbf{1}$ gives the bound $p^\star \geq n \lambda_{\min}(\mathbf{W})$.

Relation to the Conjugate Function

Consider the following optimization problem with affine inequality and equality constraints.

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b}\,,\ \mathbf{C}\mathbf{x} = \mathbf{d} \end{array} .$$

The dual function corresponding to this problem can be elegantly expressed in terms of the conjugate $f_0^{\circledast}$ of the function f_0, given by

$$f_0^{\circledast}(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{dom}(f_0)} \left\{ \mathbf{y}^T \mathbf{x} - f_0(\mathbf{x}) \right\} .$$

Dual function:

$$\begin{aligned} g(\boldsymbol{\lambda}, \boldsymbol{\nu}) &= \inf_{\mathbf{x} \in \mathrm{dom}(f_0)} \left\{ f_0(\mathbf{x}) + \left(\mathbf{A}^T \boldsymbol{\lambda} + \mathbf{C}^T \boldsymbol{\nu} \right)^T \mathbf{x} - \mathbf{b}^T \boldsymbol{\lambda} - \mathbf{d}^T \boldsymbol{\nu} \right\} , \\ &= -f_0^{\circledast}\left(-\mathbf{A}^T \boldsymbol{\lambda} - \mathbf{C}^T \boldsymbol{\nu} \right) - \mathbf{b}^T \boldsymbol{\lambda} - \mathbf{d}^T \boldsymbol{\nu} . \end{aligned}$$

This simplifies the derivation of the dual if the conjugate of f_0 is known.

Example: *Entropy maximization*

$$f_0(\mathbf{x}) = \sum_{i=1}^{n} x_i \log x_i \,,\ f_0^{\circledast}(\mathbf{y}) = \sum_{i=1}^{n} e^{y_i - 1} .$$

The Lagrange Dual Problem

The problem of finding the best lower bound on $p^\star$ obtained from the Lagrange dual function is called the *Lagrange dual problem* for the original or *primal problem*.

Lagrange Dual Problem

$$\begin{array}{ll} \text{maximize} & g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \\ \text{subject to} & \boldsymbol{\lambda} \succeq \mathbf{0} \end{array} .$$

Here, the variables are $\boldsymbol{\lambda} \in \mathbb{R}^m$ and $\boldsymbol{\nu} \in \mathbb{R}^p$.

- As this is a concave maximization problem over a convex set, this is a convex optimization problem, whose optimal value is denoted as $d^\star$.
- The variable pair $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ is said to be *dual feasible* if $\boldsymbol{\lambda} \succeq \mathbf{0}$ and $(\boldsymbol{\lambda}, \boldsymbol{\nu}) \in \mathrm{dom}(g)$.
- This is often simplified by making the implicit constraint that $(\boldsymbol{\lambda}, \boldsymbol{\nu}) \in \mathrm{dom}(g)$ explicit. Such an equivalent problem is also called the Lagrange dual problem or dual problem, with some abuse of terminology.

Example: *Standard form LP and its dual*

Primal:
$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b}\,,\ \mathbf{x} \succeq \mathbf{0} \end{array}$$

Dual:
$$\begin{array}{ll} \text{maximize} & -\mathbf{b}^T\boldsymbol{\nu} \\ \text{subject to} & \mathbf{A}^T\boldsymbol{\nu} + \mathbf{c} \succeq \mathbf{0} \end{array} .$$

Weak and Strong Duality

Weak Duality: $d^\star \leq p^\star$

- This condition always holds for both convex and nonconvex problems.
- It can be used to find nontrivial lower bounds for difficult problems.

  For example, solving the SDP

$$\begin{array}{ll} \text{maximize} & -\mathbf{1}^T \boldsymbol{\nu} \\ \text{subject to} & \mathbf{W} + \mathrm{diag}(\boldsymbol{\nu}) \succeq \mathbf{0} \end{array},$$

  gives a lower bound for the two-way partitioning problem from above.

Strong Duality: $d^\star = p^\star$

- This condition does not hold in general.
- It *usually* holds for convex problems.
- Conditions that guarantee strong duality in convex problems are called *constraint qualifications*.

Slater's Constraint Qualification

Strong duality holds for a convex optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, m \quad , \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array}$$

if it is *strictly feasible*, meaning that

$$\exists\, \mathbf{x} \in \operatorname{relint}(\mathcal{D}) : f_i(\mathbf{x}) < 0\,,\ i = 1, \ldots, m\,,\ \mathbf{A}\mathbf{x} = \mathbf{b}\,.$$

This constraint qualification is called *Slater's condition*.

- If $p^\star > -\infty$, this condition also guarantees that the dual optimal is attained.
- It can be sharpened when some of the inequality constraint functions f_i are affine. Specifically, if the first k constraint functions $f_1, \ldots, f_k$ are affine, then strong duality holds provided the following weaker condition holds.

$$\exists\, \mathbf{x} \in \operatorname{relint}(\mathcal{D}) : f_i(\mathbf{x}) \leq 0\,,\ i = 1, \ldots, k\,,\ f_i(\mathbf{x}) < 0\,,\ i = k+1, \ldots, m\,,\ \mathbf{A}\mathbf{x} = \mathbf{b}\,.$$

- There exist many other types of constraint qualifications.

Complementary Slackness

Assume that strong duality holds (i.e., $p^\star = d^\star$), $\mathbf{x}^\star$ is primal optimal, and $(\boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$ is dual optimal. Then, we have the following.

$$\begin{aligned} f_0(\mathbf{x}^\star) &= g(\boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star) = \inf_{\mathbf{x} \in \mathcal{D}} L(\mathbf{x}, \boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star) \\ &= \inf_{\mathbf{x} \in \mathcal{D}} \left\{ f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i^\star f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i^\star h_i(\mathbf{x}) \right\} , \\ &\leq f_0(\mathbf{x}^\star) + \sum_{i=1}^{m} \lambda_i^\star f_i(\mathbf{x}^\star) + \sum_{i=1}^{p} \nu_i^\star h_i(\mathbf{x}^\star) , \\ &\leq f_0(\mathbf{x}^\star) . \end{aligned}$$

Thus, the two inequalities hold with equality. From this, we ascertain the following.

- The primal optimal point $\mathbf{x}^\star$ minimizes $L(\mathbf{x}, \boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$. Note that the Lagrangian $L(\mathbf{x}, \boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star)$ may have other minimizers besides $\mathbf{x}^\star$.
- We have $\lambda_i^\star f_i(\mathbf{x}^\star) = 0$ for $i = 1, \ldots, m$. This condition is known as *complementary slackness*. Specifically, this implies that

  $$\lambda_i^\star > 0 \Longrightarrow f_i(\mathbf{x}^\star) = 0 \, , \; f_i(\mathbf{x}^\star) < 0 \Longrightarrow \lambda_i^\star = 0 \, .$$

  In other words, the slackness sparsity patterns for the primal inequality constraints and corresponding dual Lagrange multipliers do not overlap.

Inequality Form LP

Primal problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \preceq \mathbf{b} \end{array} .$$

Dual function:

$$g(\boldsymbol{\lambda}) = \inf_{\mathbf{x}} \left\{ \left(\mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda}\right)^T \mathbf{x} - \mathbf{b}^T\boldsymbol{\lambda} \right\} = \begin{cases} -\mathbf{b}^T\boldsymbol{\lambda}\,, & \mathbf{A}^T\boldsymbol{\lambda} + \mathbf{c} = \mathbf{0} \\ -\infty\,, & \text{otherwise} \end{cases} .$$

Dual problem:

$$\begin{array}{ll} \text{maximize} & -\mathbf{b}^T\boldsymbol{\lambda} \\ \text{subject to} & \mathbf{A}^T\boldsymbol{\lambda} + \mathbf{c} = \mathbf{0}\,,\ \boldsymbol{\lambda} \succeq \mathbf{0} \end{array} .$$

- From Slater's condition, $p^\star = d^\star$ if $\mathbf{A}\widetilde{\mathbf{x}} \prec \mathbf{b}$ for some $\widetilde{\mathbf{x}}$.
- As all constraints are affine, we have in fact $p^\star = d^\star$ except when both the primal and dual are infeasible.

General and Special QCQPs

Primal problem:

$$\begin{array}{ll} \text{minimize} & (1/2)\,\mathbf{x}^T\mathbf{P}_0\mathbf{x} + \mathbf{q}_0^T\mathbf{x} + r_0 \\ \text{subject to} & (1/2)\,\mathbf{x}^T\mathbf{P}_i\mathbf{x} + \mathbf{q}_i^T\mathbf{x} + r_i \le 0\,,\ i = 1,\dots,m \end{array} .$$

Here, $\mathbf{P}_0 \in \mathbb{S}^n_{++}$ and $\mathbf{P}_i \in \mathbb{S}^n_{+}$ for $i = 1, \dots, n$.

Dual function:

$$L(\mathbf{x}, \boldsymbol{\lambda}) = (1/2)\,\mathbf{x}^T\mathbf{P}(\boldsymbol{\lambda})\,\mathbf{x} + \mathbf{q}(\boldsymbol{\lambda})^T\,\mathbf{x} + r(\boldsymbol{\lambda})\ ,$$

where

$$\mathbf{P}(\boldsymbol{\lambda}) = \mathbf{P}_0 + \sum_{i=1}^{m} \lambda_i \mathbf{P}_i\,,\ \mathbf{q}(\boldsymbol{\lambda}) = \mathbf{q}_0 + \sum_{i=1}^{m} \lambda_i \mathbf{q}_i\,,\ r(\boldsymbol{\lambda}) = r_0 + \sum_{i=1}^{m} \lambda_i r_i\,.$$

If $\boldsymbol{\lambda} \succeq \mathbf{0}$, then $\mathbf{P}(\boldsymbol{\lambda}) \succ \mathbf{0}$ and we have

$$g(\boldsymbol{\lambda}) = \inf_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}) = -\,(1/2)\,\mathbf{q}(\boldsymbol{\lambda})^T\,\mathbf{P}(\boldsymbol{\lambda})^{-1}\,\mathbf{q}(\boldsymbol{\lambda}) + r(\boldsymbol{\lambda})\ .$$

Dual problem:

$$\begin{array}{ll} \text{maximize} & -\,(1/2)\,\mathbf{q}(\boldsymbol{\lambda})^T\,\mathbf{P}(\boldsymbol{\lambda})^{-1}\,\mathbf{q}(\boldsymbol{\lambda}) + r(\boldsymbol{\lambda}) \\ \text{subject to} & \boldsymbol{\lambda} \succeq \mathbf{0} \end{array} .$$

- From Slater's condition, $p^\star = d^\star$ if there exists some $\widetilde{\mathbf{x}}$ such that

$$(1/2)\,\widetilde{\mathbf{x}}^T\mathbf{P}_i\widetilde{\mathbf{x}} + \mathbf{q}_i^T\widetilde{\mathbf{x}} + r_i < 0\,,\ i = 1,\dots,m\,.$$

- When $\mathbf{P}_i = \mathbf{0}$ for all i, then all constraints are affine and so $p^\star = d^\star$ always in this case.

Nonconvex QP with Strong Duality

Primal problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{x}^T \mathbf{A} \mathbf{x} + 2\mathbf{b}^T \mathbf{x} \\ \text{subject to} & \mathbf{x}^T \mathbf{x} \leq 1 \end{array} .$$

Here, $\mathbf{A} \in \mathbb{S}^n$ but $\mathbf{A} \not\succeq \mathbf{0}$, and so the problem is not convex.

Dual function:

$$L(\mathbf{x}, \lambda) = \mathbf{x}^T \left(\mathbf{A} + \lambda \mathbf{I} \right) \mathbf{x} + 2\mathbf{b}^T \mathbf{x} - \lambda \,.$$

- The Lagrangian $L(\mathbf{x}, \lambda)$ is unbounded below if $\mathbf{A} + \lambda \mathbf{I} \not\succeq \mathbf{0}$ or if $\mathbf{A} + \lambda \mathbf{I} \succeq \mathbf{0}$ and $\mathbf{b} \notin \mathcal{R}(\mathbf{A} + \lambda \mathbf{I})$.
- It is minimized by $\mathbf{x} = -\left(\mathbf{A} + \lambda \mathbf{I} \right)^{\#} \mathbf{b}$ otherwise, in which case $g(\lambda) = -\mathbf{b}^T \left(\mathbf{A} + \lambda \mathbf{I} \right)^{\#} \mathbf{b} - \lambda$.

$$g(\lambda) = \begin{cases} -\mathbf{b}^T \left(\mathbf{A} + \lambda \mathbf{I} \right)^{\#} \mathbf{b} - \lambda \,, & \mathbf{A} + \lambda \mathbf{I} \succeq \mathbf{0} \,,\ \mathbf{b} \in \mathcal{R}(\mathbf{A} + \lambda \mathbf{I}) \\ -\infty \,, & \text{otherwise} \end{cases} .$$

Dual problem: (and equivalent SDP)

$$\begin{array}{ll} \text{maximize} & -\mathbf{b}^T \left(\mathbf{A} + \lambda \mathbf{I} \right)^{\#} \mathbf{b} - \lambda \\ \text{subject to} & \mathbf{A} + \lambda \mathbf{I} \succeq \mathbf{0} \,,\ \mathbf{b} \in \mathcal{R}(\mathbf{A} + \lambda \mathbf{I}) \end{array} , \qquad \begin{array}{ll} \text{maximize} & -t - \lambda \\ \text{subject to} & \begin{bmatrix} \mathbf{A} + \lambda \mathbf{I} & \mathbf{b} \\ \mathbf{b}^T & t \end{bmatrix} \succeq \mathbf{0} \end{array} .$$

- Strong duality (i.e., $p^\star = d^\star$) holds even though the primal problem is not convex. The proof of this involves a theorem of alternatives known as the *S-procedure*.

Weak/Strong Duality Via the Set of Values

A geometric interpretation of the dual function can be given in terms of the set of values

$$\mathcal{G} = \{(f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}), h_1(\mathbf{x}), \ldots, h_p(\mathbf{x}), f_0(\mathbf{x})) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} : \mathbf{x} \in \mathcal{D}\} .$$

With this set, the optimal value of the primal problem $p^\star$ can be expressed as

$$p^\star = \inf \{t : (\mathbf{u}, \mathbf{v}, t) \in \mathcal{G}, \ \mathbf{u} \preceq \mathbf{0}, \ \mathbf{v} = \mathbf{0}\} ,$$

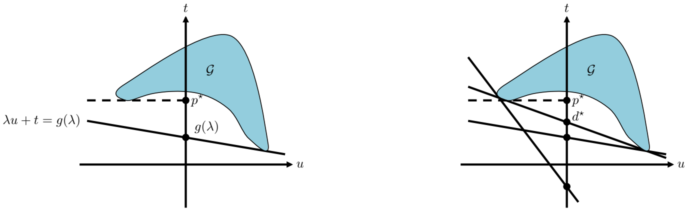
while the dual function $g(\boldsymbol{\lambda}, \boldsymbol{\nu})$ is given by

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf \left\{ (\boldsymbol{\lambda}, \boldsymbol{\nu}, 1)^T (\mathbf{u}, \mathbf{v}, t) : (\mathbf{u}, \mathbf{v}, t) \in \mathcal{G} \right\} .$$

Hence, $(\boldsymbol{\lambda}, \boldsymbol{\nu}, 1)^T (\mathbf{u}, \mathbf{v}, t) \geq g(\boldsymbol{\lambda}, \boldsymbol{\nu})$ defines a *nonvertical supporting hyperplane* to $\mathcal{G}$.

Visual example: For simplicity, consider a problem with one constraint. In this case, we have

$$g(\lambda) = \inf_{(u,t) \in \mathcal{G}} \{\lambda u + t\} , \text{ where } \mathcal{G} = \{(f_1(\mathbf{x}), f_0(\mathbf{x})) : \mathbf{x} \in \mathcal{D}\} .$$



- Here, $\lambda u + t = g(\lambda)$ is a nonvertical supporting hyperplane to $\mathcal{G}$.
- The hyperplane intersects the t-axis at $t = g(\lambda)$.

Epigraph Variation

An alternate geometric interpretation can be made by considering an *epigraph form* of $\mathcal{G}$ given by

$$\begin{aligned}\mathcal{A} &= \mathcal{G} + \left\{\mathbb{R}_+^m \times \{\mathbf{0}\} \times \mathbb{R}_+\right\} , \\ &= \{(\mathbf{u}, \mathbf{v}, t) : \exists\, \mathbf{x} \in \mathcal{D}, f_i(\mathbf{x}) \leq u_i\ ,\ i = 1, \ldots, m\ ,\ h_i(\mathbf{x}) = v_i\ ,\ i = 1, \ldots, p\ ,\ f_0(\mathbf{x}) \leq t\} \ .\end{aligned}$$

With this set, the optimal value of the primal problem $p^\star$ and the dual function $g(\boldsymbol{\lambda}, \boldsymbol{\nu})$ are given by

$$p^\star = \inf\left\{t : (\mathbf{0}, \mathbf{0}, t) \in \mathcal{A}\right\}\ ,\ g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf\left\{(\boldsymbol{\lambda}, \boldsymbol{\nu}, 1)^T (\mathbf{u}, \mathbf{v}, t) : (\mathbf{u}, \mathbf{v}, t) \in \mathcal{A}\right\} \ .$$

Thus, $(\boldsymbol{\lambda}, \boldsymbol{\nu}, 1)^T (\mathbf{u}, \mathbf{v}, t) \geq g(\boldsymbol{\lambda}, \boldsymbol{\nu})$ defines a *nonvertical supporting hyperplane* to $\mathcal{A}$.

Strong duality:

- This condition holds if and only if there exists a nonvertical supporting hyperplane to $\mathcal{A}$ at its boundary point $(\mathbf{0}, \mathbf{0}, p^\star)$.
- For a convex problem, $\mathcal{A}$ is convex, and so has a supporting hyperplane at $(\mathbf{0}, \mathbf{0}, p^\star)$.
- Slater's condition: If there exists $(\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}}, t) \in \mathcal{A}$ with $\widetilde{\mathbf{u}} \prec \mathbf{0}$ and $\widetilde{\mathbf{v}} = \mathbf{0}$, then the supporting hyperplanes at $(\mathbf{0}, \mathbf{0}, p^\star)$ must be nonvertical.

Visual example: $\mathcal{A} = \{(u, t) : f_1(\mathbf{x}) \leq u\ ,\ f_0(\mathbf{x}) \leq t \text{ for some } \mathbf{x} \in \mathcal{D}\}$