

Lagrange Duality

Yuanming Shi

ShanghaiTech University

Outline

- 1 Lagrangian
- 2 Dual Function
- 3 Dual Problem
- 4 Weak and Strong Duality
- 5 KKT conditions

Lagrangian

- Consider an optimization problem in standard form (not necessarily convex)

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

with variable $\mathbf{x} \in \mathbb{R}^n$, domain \mathcal{D} , and optimal value p^*

- The *Lagrangian* is a function $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$, with $\text{dom } L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$, defined as

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

where λ_i is the Lagrange multiplier associated with $f_i(\mathbf{x}) \leq 0$ and ν_i is the Lagrange multiplier associated with $h_i(\mathbf{x}) = 0$.

Outline

- 1 Lagrangian
- 2 Dual Function**
- 3 Dual Problem
- 4 Weak and Strong Duality
- 5 KKT conditions

Lagrange Dual Function I

- The *Lagrange dual function* is defined as the infimum of the Lagrangian over $\mathbf{x} : g : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$,

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

- Observe that:
 - the infimum is unconstrained (as opposed to the original constrained minimization problem)
 - g is concave regardless of original problem (infimum of affine functions)
 - g can be $-\infty$ for some $\boldsymbol{\lambda}, \boldsymbol{\nu}$

Lagrange Dual Function II

🦉 **Lower bound property:** if $\lambda \succeq 0$, then $g(\lambda, \nu) \leq p^*$.

Proof.

Suppose \tilde{x} is feasible and $\lambda \succeq 0$. Then,

$$f_0(\tilde{x}) \geq L(\tilde{x}, \lambda, \nu) \geq \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = g(\lambda, \nu)$$

Now choose minimizer of $f_0(\tilde{x})$ over all feasible \tilde{x} to get $p^* \geq g(\lambda, \nu)$. \square

🦉 We could try to find the best lower bound by maximizing $g(\lambda, \nu)$.
This is in fact the dual problem.

Outline

- 1 Lagrangian
- 2 Dual Function
- 3 Dual Problem**
- 4 Weak and Strong Duality
- 5 KKT conditions

Dual Problem

- The *Lagrange dual problem* is defined as

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

- This problem finds the best lower bound on p^* obtained from the dual function
- It is a convex optimization (maximization of a concave function and linear constraints)
- The optimal value is denoted d^*
- λ, ν are dual feasible if $\lambda \succeq \mathbf{0}$ and $(\lambda, \nu) \in \text{dom } g$ (the latter implicit constraints can be made explicit in problem formulation)

Example: Least-Norm Solution of Linear Equations I

- Consider the problem

$$\begin{array}{ll}\underset{x}{\text{minimize}} & x^T x \\ \text{subject to} & Ax = b\end{array}$$

- The Lagrangian is

$$L(x, \nu) = x^T x + \nu^T (Ax - b)$$

- To find the dual function, we need to solve an unconstrained minimization of the Lagrangian. We set the gradient equal to zero

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \implies x = -\frac{1}{2} A^T \nu$$

Example: Least-Norm Solution of Linear Equations II

and we plug the solution in L to obtain g :

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

- The function g is, as expected, a concave function of $\boldsymbol{\nu}$.
- From the lower bound property, we have

$$p^* \geq -\frac{1}{4}\boldsymbol{\nu}^T\mathbf{A}\mathbf{A}^T\boldsymbol{\nu} - \mathbf{b}^T\boldsymbol{\nu} \text{ for all } \boldsymbol{\nu}$$

- The dual problem is the QP

$$\underset{\boldsymbol{\nu}}{\text{maximize}} \quad -\frac{1}{4}\boldsymbol{\nu}^T\mathbf{A}\mathbf{A}^T\boldsymbol{\nu} - \mathbf{b}^T\boldsymbol{\nu}$$

Example: Standard Form LP I

- Consider the problem

$$\begin{array}{ll}\underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & Ax = b, \quad x \succeq 0\end{array}$$

- The Lagrangian is

$$\begin{aligned}L(x, \lambda, \nu) &= c^T x + \nu^T (Ax - b) - \lambda^T x \\ &= (c + A^T \nu - \lambda)^T x - b^T \nu\end{aligned}$$

- L is a linear function of x and it is unbounded if the term multiplying x is nonzero.

Example: Standard Form LP II

☞ Hence, the dual function is

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

☞ The function g is a concave function of (λ, ν) as it is linear on an affine domain.

☞ From the lower bound property, we have

$$p^* \geq -b^T \nu \quad \text{if } c + A^T \nu \succeq 0$$

☞ The dual problem is the LP

$$\begin{array}{ll} \underset{\nu}{\text{maximize}} & -b^T \nu \\ \text{subject to} & c + A^T \nu \succeq 0 \end{array}$$

Outline

- 1 Lagrangian
- 2 Dual Function
- 3 Dual Problem
- 4 Weak and Strong Duality**
- 5 KKT conditions

Weak and Strong Duality I

- From the lower bound property, we know that $g(\lambda, \nu) \leq p^*$ for feasible (λ, ν) . In particular, for a (λ, ν) that solves the dual problem.
- Hence, **weak duality** always holds (even for nonconvex problems):

$$d^* \leq p^*$$

- The difference $p^* - d^*$ is called **duality gap**.
- Solving the dual problem may be used to find nontrivial lower bounds for difficult problems.
- Even more interesting is when equality is achieved in weak duality. This is called **strong duality**:

$$d^* = p^*$$

Weak and Strong Duality II

- Strong duality means that the duality gap is zero.
- Strong duality:
 - is very desirable (we can solve a difficult problem by solving the dual)
 - does not hold in general
 - usually holds for convex problems
 - conditions that guarantee strong duality in convex problems are called **constraint qualifications**.

Slater's Constraint Qualification I

- Slater's constraint qualification is a very simple condition that is satisfied in most cases and ensures strong duality for convex problems.
- Strong duality holds for a convex problem

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

if it is strictly feasible, i.e.,

$$\exists \mathbf{x} \in \text{int } \mathcal{D} : \quad f_i(\mathbf{x}) < 0 \quad i = 1, \dots, m, \quad \mathbf{Ax} = \mathbf{b}$$

- There exist many other types of constraint qualifications.

Example: Inequality Form LP

- Consider the problem

$$\begin{array}{ll}\underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & Ax \preceq b\end{array}$$

- The dual problem is

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

- From Slater's condition: $p^* = d^*$ if $A\tilde{x} \prec b$ for some \tilde{x} .
- In this case, in fact, $p^* = d^*$ except when primal and dual are infeasible.

Example: Convex QP

- Consider the problem (assume $P \succeq 0$)

$$\begin{array}{ll}\underset{x}{\text{minimize}} & x^T P x \\ \text{subject to} & Ax \preceq b\end{array}$$

- The dual problem is

$$\begin{array}{ll}\underset{\lambda}{\text{maximize}} & -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda \\ \text{subject to} & \lambda \succeq 0\end{array}$$

- From Slater's condition: $p^* = d^*$ if $A\tilde{x} \prec b$ for some \tilde{x} .
- In this case, in fact, $p^* = d^*$ always.

Complementary Slackness

- Assume strong duality holds, \mathbf{x}^* is primal optimal and $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ is dual optimal. Then

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

- Hence, the two inequalities must hold with equality. Implications:
 - \mathbf{x}^* minimizes $L(\mathbf{x}, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$
 - $\lambda_i^* f_i(\mathbf{x}^*) = 0$ for $i = 1, \dots, m$; this is called **complementary slackness**:

$$\lambda_i^* > 0 \implies f_i(\mathbf{x}^*) = 0, \quad f_i(\mathbf{x}^*) < 0 \implies \lambda_i^* = 0$$

Outline

- 1 Lagrangian
- 2 Dual Function
- 3 Dual Problem
- 4 Weak and Strong Duality
- 5 KKT conditions**

Karush-Kuhn-Tucker (KKT) Conditions

KKT conditions (for differentiable f_i, h_i):

1 primal feasibility:

$$f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \quad h_i(\mathbf{x}) = 0, \quad i = 1, \dots, p$$

2 dual feasibility: $\lambda \succeq \mathbf{0}$

3 complementary slackness: $\lambda_i f_i(\mathbf{x}) = 0$ for $i = 1, \dots, m$

4 zero gradient of Lagrangian with respect to \mathbf{x} :

$$\nabla f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla f_i(\mathbf{x}) + \sum_{i=1}^p \nu_i \nabla h_i(\mathbf{x}) = \mathbf{0}$$

KKT condition

- We already know that if strong duality holds and x, λ, ν are optimal, then they must satisfy the KKT conditions.
- What about the opposite statement?
- If x, λ, ν satisfy the KKT conditions for a convex problem, then they are optimal.

Proof.

From complementary slackness, $f_0(x) = L(x, \lambda, \nu)$ and, from 4th KKT condition and convexity, $g(\lambda, \nu) = L(x, \lambda, \nu)$. Hence, $f_0(x) = g(\lambda, \nu)$. \square

Theorem

If a problem is convex and Slater's condition is satisfied, then x is optimal if and only if there exists λ, ν that satisfy the KKT conditions.

Reference

Chapter 5 of:

- Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.