Lecture 20: Duality of Nonlinear Optimization

Zizhuo Wang

Institute of Data and Decision Analytics (iDDA)
Chinese University of Hong Kong, Shenzhen

Nov 16, 2018

Announcements

- Homework 7 posted, due next Wednesday (11/21)

Recap: Convexity/Convex Optimization Problem

We call an optimization problem a convex optimization problem if

- The objective is either minimizing a convex function or maximizing a concave function
- Feasible set is a convex set

A sufficient condition for the feasible set to be a convex set is that each constraint is a convex constraint.

Sufficient conditions for a constraint to be a convex constraint:

- $f(\mathbf{x}) \leq 0$ with $f(\mathbf{x})$ being convex
- $g(\mathbf{x}) \geq 0$ with $g(\mathbf{x})$ being concave
- Any linear constraint

Sometimes, the constraints are not in the right form, but one can make a transformation to make it a convex constraint.

Verifying Constraints

Often one makes monotone transformation or applies variable substitutions. But sometimes one has to look into the region the constraints define (using definition):

- $x^3 - 1 \leq 0$: x^3 is not a convex function, however, this constraint defines a convex feasible region ($x \leq 1$) thus is a convex constraint.
- $z^2 - xy \leq 0$, $x, y, z \geq 0$: $z^2 - xy$ is not a convex function. The Hessian matrix is

  $$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

  The eigenvalues are -1, 1 and 2
- However, this gives a convex region

Proof

We prove by definition. We show for any (x_1, y_1, z_1) and (x_2, y_2, z_2) that satisfy these constraints, and for any $0 \leq \alpha \leq 1$,

$$(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2, \alpha z_1 + (1-\alpha)z_2)$$

also satisfy this constraint (i.e., the feasible region is convex)

We have

$$(\alpha z_1 + (1-\alpha)z_2)^2 = \alpha^2 z_1^2 + (1-\alpha)^2 z_2^2 + 2\alpha(1-\alpha)z_1 z_2$$

By assumption, $z_1^2 \leq x_1 y_1$, $z_2^2 \leq x_2 y_2$. Also when $x, y, z \geq 0$, we have

$$2z_1 z_2 \leq 2\sqrt{x_1 y_1 x_2 y_2} \leq x_1 y_2 + x_2 y_1$$

The last inequality is because of the inequality of arithmetic and geometric means

Proof Continued

Therefore

$$\begin{aligned}
&(\alpha z_1 + (1-\alpha) z_2)^2 \\
= &\alpha^2 z_1^2 + (1-\alpha)^2 z_2^2 + 2\alpha(1-\alpha) z_1 z_2 \\
\le &\alpha^2 x_1 y_1 + (1-\alpha)^2 x_2 y_2 + \alpha(1-\alpha)(x_1 y_2 + x_2 y_1) \\
= &(\alpha x_1 + (1-\alpha) x_2)(\alpha y_1 + (1-\alpha) y_2)
\end{aligned}$$

Thus $z^2 - xy \le 0,\ x, y, z \ge 0$ is a convex region. $\square$

We don't often use this method, but this can be viewed as a last resort.

Software to Solve Nonlinear Optimization Problems

MATLAB has its own functions: fmincon and fminunc.

- However, in general they are not very scalable nor fast

We suggest to use CVX. CVX can solve a range of nonlinear optimization problems.

- CVX can only solve convex optimization problems (that is what it is named for). It can only recognize certain classes of convex functions
- Sometimes, one has to manually convert a problem into a recognizable form before inputting into CVX

Examples

Example 1:

$$\begin{array}{ll} \text{minimize} & (x_1 - 1)^2 + (x_2 - 1)^2 \\ \text{s.t.} & x_1 + x_2 = 1 \end{array}$$

Example 2:

$$\begin{array}{ll} \text{minimize} & e^{x_1 + x_2} + (x_1 - 0.5x_2)^2 + 2.75x_2^2 \\ \text{s.t.} & x_1 + 2x_2 = 1 \end{array}$$

Examples

Least Squares Problem

$$\begin{array}{rl} \text{minimize}_{\beta} & ||X\beta - \mathbf{y}||_2^2 \\ \text{s.t.} & W\beta = \xi \end{array}$$

The KKT conditions:

$$\left[\begin{array}{c} \beta \\ \mathbf{z} \end{array}\right] = \left[\begin{array}{cc} W & 0 \\ X^T X & -\frac{1}{2} W^T \end{array}\right]^{-1} \left[\begin{array}{c} \xi \\ X^T \mathbf{y} \end{array}\right]$$

Importance of Convexity

Theorem

For convex optimization problems, all local minimizers must be global minimizers.

Theorem

For convex optimization problems, the KKT conditions are sufficient for global optimality.

Therefore, convex optimization problems have uniquely good property, and is the main criterion to determine whether a problem is easy or hard to solve.

Agenda

- Discuss duality in nonlinear case

Duality for General Optimization Problems

Consider a minimization problem (F contains all the sign constraints on $\mathbf{x}$):

$$\begin{array}{rl} \text{minimize}_{\mathbf{x}\in F} & f(\mathbf{x}) \\ \text{subject to} & g_i(\mathbf{x}) \geq 0, \quad \forall i \\ & h_i(\mathbf{x}) = 0, \quad \forall i \\ & \ell_i(\mathbf{x}) \leq 0, \quad \forall i \end{array}$$

Remember when deriving the KKT conditions, we first construct the Lagrangian:

$$L(\mathbf{x}, \lambda, \nu, \eta) = f(\mathbf{x}) + \sum_i \lambda_i g_i(\mathbf{x}) + \sum_i \nu_i h_i(\mathbf{x}) + \sum_i \eta_i \ell_i(\mathbf{x})$$

We also have the *dual feasibility conditions* $\lambda_i \leq 0$ and $\eta_i \geq 0$.

Duality for General Optimization Problems

We claim the original problem

$$\begin{array}{rll} \text{minimize}_{\mathbf{x}\in F} & f(\mathbf{x}) & \\ \text{subject to} & g_i(\mathbf{x}) \geq 0, & \forall i \\ & h_i(\mathbf{x}) = 0, & \forall i \\ & \ell_i(\mathbf{x}) \leq 0 & \forall i \end{array}$$

is equivalent to:

$$\text{minimize}_{\mathbf{x}\in F} \quad \{\text{maximize}_{\lambda\leq 0,\nu,\eta\geq 0} \quad L(\mathbf{x},\lambda,\nu,\eta)\}$$

Why? In order for the inner problem to be bounded, we must have $g_i(\mathbf{x}) \geq 0$, $h_i(\mathbf{x}) = 0$ and $\ell_i(\mathbf{x}) \leq 0$ and in this case

$$\max_{\lambda\leq 0,\nu,\eta\geq 0} L(\mathbf{x},\lambda,\nu,\eta) = f(\mathbf{x})$$

This reminds us about when we derived the dual problem for LP.

Duality for General Optimization Problems

Our original problem (primal problem) is equivalent to:

$$\min_{\mathbf{x}\in F} \quad \{\max_{\lambda\le 0,\nu,\eta\ge 0} L(\mathbf{x},\lambda,\nu,\eta)\}$$

Recall for LP, we exchanged the min and max and derived the dual problem. Here we do the same:

$$\max_{\lambda\le 0,\nu,\eta\ge 0} \quad \{\min_{\mathbf{x}\in F} L(\mathbf{x},\lambda,\nu,\eta)\}$$

We call this problem the dual problem.

- That is why $\lambda \le 0$, $\eta \ge 0$ are called *dual feasibility* constraints

Duality for General Optimization Problems

Now we have the dual problem (again we did not justify the exchange of min and max):

$$\max_{\lambda \leq 0, \nu, \eta \geq 0} \quad \{\min_{\mathbf{x} \in F} L(\mathbf{x}, \lambda, \nu, \eta)\}$$

- λ, ν and η are dual variables
- If the inner problem has an explicit form, then the dual problem is an explicit maximization problem for λ, ν and η (this was the case for linear optimization problems)
- However, in general, the inner problem may not have an explicit solution. In those cases, the dual problem doesn't have an explicit form (but it is still an optimization problem for λ, ν and η)

Example

Consider the following problem:

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

It is a nonlinear optimization problem. Now let's construct its dual problem. We start from the Lagrangian:

$$L(\mathbf{x}, \lambda) = \mathbf{x}^T\mathbf{x} + \lambda^T(A\mathbf{x} - \mathbf{b})$$

with dual feasibility constraints $\lambda \geq 0$.

Example Continued

Therefore, the dual problem is:

$$\max_{\lambda \geq 0} \quad \{\min_{\mathbf{x}} \quad \mathbf{x}^T \mathbf{x} + \lambda^T A\mathbf{x} - \mathbf{b}^T \lambda\}$$

The inner problem is an unconstrained convex problem. By the KKT conditions, the optimal solution is

$$\mathbf{x}^* = -\frac{1}{2} A^T \lambda$$

Therefore, the inner problem has an explicit optimal value

$$-\frac{1}{4} \lambda^T A A^T \lambda - \mathbf{b}^T \lambda$$

Example Continued

Therefore, the dual problem of

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

is

$$\begin{array}{ll} \text{maximize}_{\lambda} & -\frac{1}{4}\lambda^T AA^T\lambda - \mathbf{b}^T\lambda \\ \text{subject to} & \lambda \ge 0 \end{array}$$

Still, the dual of a minimization is a maximization.

- However, there is no easy rule to find the dual as in the LP case, usually one has to go through the above steps to find the dual (and usually get some implicit form)

Duality Theorems

Recall we have weak duality theorem and strong duality theorem for LPs. Now we still have the weak duality theorem.

Theorem (Weak Duality Theorem)

Given any minimization problem and its dual. Then

- *The objective value of any feasible solution for the primal problem is an upper bound on the dual optimal value;*
- *The objective value of any feasible solution for the dual problem is a lower bound on the primal optimal value;*
- *The optimal value of the primal problem is always greater than or equal to the optimal value of the dual problem*
- *If one is unbounded, then the other one must be infeasible.*

Strong Duality Theorem

For LP, we have the strong duality theorem, saying that the optimal values of primal and dual problems are always the same.

Unfortunately, this is not necessarily true for nonlinear optimization problems. We call the difference between the primal optimal value and dual optimal value the *duality gap*

- For LP, the duality gap is always 0
- For general optimization problems, there could be a positive duality gap
- A positive duality gap means the exchange of max and min does not maintain the value
- Again, by the weak duality theorem, it must be that the optimal value of the minimization problem is larger than that of the maximization problem.

Strong Duality Theorem

However, there are some cases that strong duality theorem holds.

Theorem (Strong Duality Theorem)

If the primal problem is a convex optimization problem and also Slater's condition holds, then the duality gap is 0.

Definition (Slater's Condition)

Assume the primal constraints are $f_i(\mathbf{x}) \le 0$, for all i and some $A\mathbf{x} \le \mathbf{b}$. Then if there exists a feasible solution $\mathbf{x}$ such that $f_i(\mathbf{x}) < 0$, then Slater's condition holds.

- In other words, Slater's condition requires that there is a strict interior solution for nonlinear constraints.

To Summarize

- There are dual problems for nonlinear optimization problems
- Dual problems for nonlinear optimization do not necessarily have an explicit form
- Weak duality always holds
- Strong duality holds when the problem is convex and the Slater's condition holds (otherwise, there could be a positive duality gap)