Notes 05: Dual programs

1. Convex programs [BV §4.1.1,§4.2.1]

A convex program is an optimization problem minimizing a convex objective function over a convex domain.

We will consider optimization problem of the form

$$\begin{array}{llll} \text{minimize} & f_0(x) & & \text{(objective function)} \\ \text{subject to} & f_i(x) \leqslant 0 & 1 \leqslant i \leqslant m & \text{(inequality constraints)} \\ & h_i(x) = 0 & 1 \leqslant i \leqslant p & \text{(equality constraints)} \end{array}$$

where $x \in \mathbb{R}^n$ is the optimization variable, $f_0 : \mathbb{R}^n \to \mathbb{R}$ the objective function, $f_i : \mathbb{R}^n \to \mathbb{R}$ are functions in the inequality constraints, $h_i : \mathbb{R}^n \to \mathbb{R}$ are functions in the equality constraints. A point $x \in \mathbb{R}^n$ is feasible for the problem if it satisfies all the constraints.

The optimization problem is convex if $f_0, f_1, \dots, f_m$ are all convex functions, and $h_1, \dots, h_p$ are all affine functions (of the form $h_i(x) = a_i^\top x - b_i$), in which case equality constraints reduce to $a_i^\top x = b_i$. The feasible region (set of feasible points) of a convex optimization problem is convex.

1.1. Linear programs (LP).

Linear programs are the special cases where the objective function f_0 and the functions $f_1, \dots, f_m$ in the inequality constraints are all affine. In other words,

$$\begin{array}{lll} \min & c^\top x & \\ \text{subject to} & a_i^\top x \leqslant b_i & 1 \leqslant i \leqslant m \\ & d_i^\top x = s_i & 1 \leqslant i \leqslant p \end{array}$$

The LPs we define here look different from what we defined in Lecture 01, but there are standard tricks to convert from between these two representations.

1.2. Semidefinite programs (SDP).

Semidefinite programs are special cases of convex programs, where $x \in \mathbb{R}^n$ corresponds to the upper triangular entries of a real symmetric matrix X. Again the objective function f_0 is affine, and functions $f_1, \dots, f_m$ in the inequality constraints are either affine or the negative minimum eigenvalue function $f_i(x) = -\lambda_{\min}(X)$ (which is a convex function of x).

2. Local vs global optimality [BV §4.2.2]

Unilke integer programs, convex programs can be solved in polynomial time up to arbitrary precision, thanks to two properties: (1) the feasible region of a convex program is convex, and (2) every locally optimal solution is automatically a globally optimal.

A locally optimal solution to a convex program is a point $x_0 \in \mathbb{R}^n$ such that, for some radius $r > 0$, x_0 minimizes the objective function f_0 among all feasible points z that has distance at most r from x (so that $\|z - x_0\| \leqslant r$).

To see that a locally optimal solution x_0 is globally optimal, consider any feasible point x. Let L be the line segment between x_0 and x. This line segment stay inside the feasible region, because the program is convex. x is also a local minimum for f_0 restricted to this line segment. Finally, one can show that a local minimum of the convex function f_0 on a line segment is also its global minimum, so $f_0(x_0) \leqslant f(x)$, as required.

3. Dual programs

Consider the following linear program:

$$\begin{array}{rr} \min & -2x_1 - 3x_2 \\ & -4x_1 - 8x_2 \geqslant -12 \\ & -2x_1 - x_2 \geqslant -3 \\ & x_1 \geqslant 0 \\ & x_2 \geqslant 0 \end{array}$$

To upperbound its objective value, we can show you one feasible solution of small value, such as $x_1 = x_2 = 1$ with objective value -5.

What about lowerbounding the objective value?

Multiply the first inequality constraint by 1/2, we get $-2x_1 - 4x_2 \geqslant -6$. Now add the last constraint $x_2 \geqslant 0$, we get $-2x_1 - 3x_2 \geqslant -6$.

To get a better lowerbound, we add the first two inequalities and divide by 3, showing $-2x_1 + -3x_2 \geqslant -5$. So the optimum value is -5.

We are trying to find the best nonnegative multipliers to add the inequalities to get the best possible lowerbound (nonnegative to avoid flipping the inequality sign). This is the dual program.

4. Langrangian dual [BV §5.1.1-5.1.3]

Definition 4.1. The Lagrangian for a convex problem is

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

where λ_i are the Lagrangian multipliers of the ith inequality constraints and ν_i are the Lagrangian multipliers of the ith equality constraints.

The Lagrangian dual function is the infimum of the Lagrangian over $x \in \mathbb{R}^n$:

$$g(\lambda, \nu) = \inf \left\{ f_0(x) + \sum_{1 \leqslant i \leqslant m} \lambda_i f_i(x) + \sum_{1 \leqslant i \leqslant p} \nu_i h_i(x) \;\middle|\; x \in \mathbb{R}^n \right\}$$

Lagrangian dual is closely related to conjugate, see [BV §5.1.6].

Given any $\lambda \in \mathbb{R}^m, \nu \in \mathbb{R}^p$ with $\lambda \geqslant 0$, $g(\lambda, \nu)$ is a lowerbound to $f_0(x)$ for any feasible x, because

$$g(\lambda, \nu) \leqslant f_0(x) + \sum_{1 \leqslant i \leqslant m} \underbrace{\lambda_i}_{\geqslant 0} \underbrace{f_i(x)}_{\leqslant 0} + \sum_{1 \leqslant i \leq p} \nu_i \underbrace{h_i(x)}_{=0} \leqslant f_0(x).$$

This inequality also holds when we take the infimum over all feasible x, and we take the supremum over all $\lambda \geqslant 0$:

$$d^\star := \sup \{g(\lambda, \nu) \mid \lambda \geqslant 0\} \leqslant \inf \{f_0(x) \mid \text{feasible } x\} =: p^\star.$$

The dual program is

$$\sup \{g(\lambda, \nu) \mid \lambda \geqslant 0\}.$$

The Lagrangian multipliers λ_i and ν_i are the dual variables of the dual program.

The dual optimal value $d^\star$ always lowerbounds the optimal value $p^\star$ of the primal (i.e. original program). This is known as weak duality.

Even if the primal program is not convex, the dual program is always convex. This is because $g(\lambda, \nu)$ is the pointwise supremum of concave (in fact, affine) functions, so $g(\lambda, \nu)$ is a concave function, and maximizing a concave function is a convex program.

How good is the dual optimum as a lowerbound?

For convex programs, under a mild condition (Slater's condition), the dual optimum gives the best lowerbound and equals the primal optimum.

4.1. **Linear programs [BV §5.1.5].** Linear program in inequality form

$$\begin{aligned} \min \quad & c^\top x \\ & Ax \leqslant b \end{aligned}$$

has Lagrangian dual

$$g(\lambda) = \inf_x \; c^\top x + \lambda^\top (Ax - b) = \inf_x \; (c^\top + \lambda^\top A)x - \lambda^\top b$$

The infimum is $-\infty$ if $c^\top + \lambda^\top A \neq 0$, and $-\lambda^\top b$ otherwise. The dual program is

$$\begin{aligned} \max \quad & -b^\top \lambda \\ & A\lambda + c = 0 \\ & \lambda \geqslant 0, \end{aligned}$$

itself a linear program in standard form.