05. Duality

By Yang Lin[1] (2024 秋季, @NJU)

Contents

[1] Institute: Nanjing University. Email: linyang@nju.edu.cn.

1 Lagrange dual problem

Lagrangian: standard form problem (not necessarily convex)

$$\begin{array}{ll} \texttt{minimize} & f_0(x) \\ \texttt{subject to} & f_i(x) \leq 0, \;\; i = 1, 2, \ldots, m \\ & h_i(x) = 0, \;\; i = 1, 2, \ldots, p \end{array}$$

Variable $x \in \mathbf{R}^n$, domain $\mathcal{D}$, optimal value p^*

Lagrangian: $L : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$, with $\mathbf{dom}\, L = \mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$,

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

Weighted sum of objective and constraint functions

λ_i is Lagrange multiplier associated with $f_i(x) \leq 0$

ν_i is Lagrange multiplier associated with $h_i(x) = 0$

1 Lagrange dual problem

Lagrange dual function: $g : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$,

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

g is concave, can be $-\infty$ for some λ, ν

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

proof.

if $\tilde{x}$ is feasible and $\lambda \succeq 0$, then

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

Minimizing over all feasible $\tilde{x}$ gives $p^* \geq g(\lambda, \nu)$

1 Lagrange dual problem

Least-norm solution of linear equations

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

dual function:

1. Lagrangian is $L(x, \nu) = x^T x + \nu^T (Ax - b)$
2. To minimize L over x, set gradient equal to zero:

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \Longrightarrow x = -(1/2) A^T \nu$$

Plug into L to obtain g:

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

A concave function of ν

Lower bound property: $p^* \geq -(1/4) \nu^T A A^T \nu - b^T \nu$ for all ν

1 Lagrange dual problem

Standard form LP

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

1. Lagrangian is

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

2. L is affine in x, hence

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

g is linear on affine domain $\{(\lambda,\nu) | A^T \nu - \lambda + c = 0\}$, hence concave
Lower bound property: $p^* \geq -b^T \nu$ for all ν if $A^T \nu + c \succeq 0$

1 Lagrange dual problem

Equality constrained norm minimization

$$\begin{array}{ll} \texttt{minimize} & \|x\| \\ \texttt{subject to} & Ax = b \end{array}$$

dual function:

$$g(\nu) = \inf_x (\|x\| - \nu^T Ax + b^T\nu) = \begin{cases} b^T\nu & \|A^T\nu\|_* \le 1 \\ -\infty & \text{otherwise} \end{cases}$$

where $\|\nu\|_* = \sup_{\|u\|\le 1} u^T\nu$ is dual norm of $\|\cdot\|$

proof. follows from $\inf_x(\|x\| - y^T x) = 0$ if $\|y\|_* \le 1, -\infty$, otherwise

1. if $\|y\|_* \le 1$, $\inf_{x>0}(\|x\| - y^T x) = \inf_{x>0}(1 - y^T x/\|x\|)\|x\| \ge 0$, with equality if $x = 0$
2. if $\|y\|_* > 1$, choose $x = tu$ where $\|u\| \le 1$, $u^T y = \|y\|_* > 1$:

$$\|x\| - y^T x = t(\|u\| - \|y\|_*) \to -\infty \text{ as } t \to \infty$$

Lower bound property: $p^* \ge b^T\nu$ if $\|A^T\nu\|_* \le 1$

1 Lagrange dual problem

Two-way partitioning

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

1. A non-convex problem; feasible set contains 2^n discrete points
2. interpretation: partition $\{1, \ldots, n\}$ in two sets; W_{ij} is cost of assigning i, j to the same set; $-W_{ij}$ is cost of assigning to different sets

dual function:

$$g(\nu) = \inf_x \left(x^T W x + \sum_i \nu_i (x_i^2 - 1) \right) = \inf_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu$$

$$= \begin{cases} -\mathbf{1}^T \nu, & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

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

Example: $\nu = -\lambda_{\min}(W)\mathbf{1}$ gives bound $p^* \geq n\lambda_{\min}(W)$

1 Lagrange dual problem

Lagrange dual and conjugate function

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

dual function:

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

1. Recall definition of conjugate $f^*(y) = \sup_{x \in \mathbf{dom}\ f}(y^T x - f(x))$
2. Simplifies derivation of dual if conjugate of f_0 is known

1 Lagrange dual problem

The dual problem

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

1. Finds best lower bound on p^*, obtained from Lagrange dual function
2. A convex optimization problem; optimal value denoted d^*
3. λ, ν are dual feasible if $\lambda \succeq 0$, $(\lambda, \nu) \in \mathbf{dom}\ g$
4. Often simplified by making implicit constraint $(\lambda, \nu) \in \mathbf{dom}\ g$ explicit

Example: standard form LP and its dual (page 5–5)

$$\begin{array}{ll} \texttt{minimize} & c^T x \\ \texttt{subject to} & Ax = b \\ & x \succeq 0 \end{array} \qquad \begin{array}{ll} \texttt{maximize} & -b^T \nu \\ \texttt{subject to} & A^T \nu + c \succeq 0 \end{array}$$

2 Weak and strong duality

Weak duality: $d^* \leq p^*$

1. Always holds(for convex and nonconvex problems)
2. Can be used to find nontrivial lower bounds for difficult problems for example, solving the SDP

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

Gives a lower bound for the two-way partitioning problem on page 5–7

Stong duality: $d^* = p^*$

1. does not hold in general
2. (usually) holds for convex problems
3. conditions that guarantee strong duality in convex problems are called **constraint qualifications**

2 Weak and strong duality

Slater’ s constraint qualification

Strong duality holds for a convex problem

$$\begin{array}{ll} \texttt{maximize} & -f_0(x) \\ \texttt{subject to} & f_i(x) \leq 0, \ \ i = 1, \ldots, m \\ & Ax = b \end{array}$$

if it is strictly feasible, i.e.,

$$\exists x \in \textbf{int}\ \mathcal{D}: \ f_i(x) < 0, \ \ i = 1, \ldots, m, \ \ Ax = b$$

1. also guarantees that the dual optimum is attained (if $p^* > -\infty$)

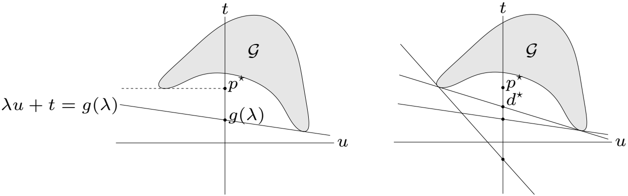
2.there exist many other types of constraint qualifications

3 Geometric interpretation

For simplicity, consider problem with one constraint $f_1(x) \leq 0$

interpretation of dual function:

$$g(\lambda) = \inf_{(u,t)\in\mathcal{G}} (t + \lambda u), \quad \mathcal{G} = \{(f_1(x), f_0(x)) | x \in \mathcal{D}\}$$

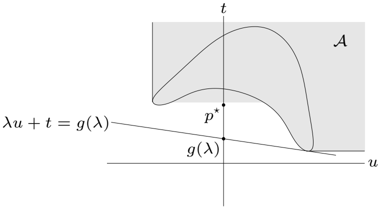


1. $\lambda u + t = g(\lambda)$ is (non-vertical) supporting hyperplane to $\mathcal{D}$
2. hyperplane intersects t-axis at $t = g(\lambda)$

3 Geometric interpretation

epigraph variation: same interpretation if $\mathcal{G}$ is replaced with

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



Strong duality

1. Holds if there is a non-vertical supporting hyperplane to $\mathcal{A}$ at $(0, p^*)$

2. For convex problem, $\mathcal{A}$ is convex, hence has supp. hyperplane at $(0, p^*)$

3. Slater's condition: if there exist $(\tilde{u}, \tilde{t}) \in \mathcal{A}$ with $\tilde{u} < 0$, then supporting hyperplanes at $(0, p^*)$ must be non-vertical

A proof of slater's condition

Assumptions: 1. contain non-empty interior point; 2. **rank** $A = p$. Define the following two exclusive sets:

$$\mathcal{A} = \{(u, v, t) | \exists x \in \mathcal{D}, f_i(x) \leq u_i, h_i(x) = v_i, f_0(x) \leq t\}$$
$$\mathcal{B} = \{(0, 0, s) \in \mathbf{R}^m \times \mathbf{R}^p \times \mathbf{R} | s < p^*\}$$

$\exists(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$ and α, such that

$$(u, v, t) \in \mathcal{A} \Longrightarrow \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \geq \alpha$$
$$(u, v, t) \in \mathcal{B} \Longrightarrow \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \leq \alpha$$

We have 1. $\tilde{\lambda} \succeq 0$ and $\mu \geq 0$ (due to feature of epigraph); 2. $\mu p^* \leq \alpha$

In other words, there exists feasible $x \in \mathcal{D}$

$$\sum \tilde{\lambda}_i f_i(x) + \tilde{\nu}^T (Ax - b) + \mu f_0(x) \geq \alpha \geq \mu p^*$$

A proof of slater's condition (cont'd)

Case 1: $\mu > 0$

$L(x, \tilde{\lambda}/\mu, \tilde{\nu}/\mu) \geq p^*$, hence, $g(\tilde{\lambda}/\mu, \tilde{\nu}/\mu) \geq p^*$

$g(\tilde{\lambda}/\mu, \tilde{\nu}/\mu) = p^*$ proved

Case 2: $\mu = 0$

$$\sum \tilde{\lambda}_i f_i(x) + \tilde{\nu}^T(Ax - b) \geq 0, \forall x \in \mathcal{D}$$

Assume $\tilde{x}$ satisfies the slater's conditions. Then $A\tilde{x} - b = 0$ and

$$\sum \tilde{\lambda}_i f_i(\tilde{x}) \geq 0 \Longrightarrow \tilde{\lambda} = 0$$

As a result, $\tilde{\nu} \neq 0$ (non-zero parameters for a hyperplane). Consider the following facts: 1. $\tilde{\nu}^T(A\tilde{x} - b) = 0$, 2. $\tilde{x} \in \textbf{int}\ \mathcal{D}$. There exists x, such that $\tilde{\nu}^T(Ax - b) < 0$ unless $A^T\tilde{v} = 0$. ($\textbf{rank} A = p$, $\tilde{\nu}$ is p-dimension).

Example: Weak and strong duality

Inequality form LP

Primal problem

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

Dual function

$$g(\lambda) = \inf_x \left((c + A^T\lambda)^T x - b^T\lambda \right) = \begin{cases} -b^T\lambda & A^T\lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases} \tag{1}$$

Dual problem

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

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

Example: Weak and strong duality

Quadratic program

Primal problem (assume $P \in \mathbf{S}^n_{++}$)

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

Dual function

$$g(\lambda) = \inf_x \left(x^T P x + \lambda^T (AX - B)\right) = -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda \quad (2)$$

Dual problem

$$\begin{array}{ll} \texttt{maximize} & -(1/4)\lambda^T A P^{-1} A^T \lambda - b^T \lambda \\ \texttt{subject to} & \lambda \succeq 0 \end{array}$$

1. from Slater' s condition: $p^* = d^*$ if $A\tilde{x} \prec b$ for some $\tilde{x}$
2. in fact, $p^* = d^*$ always holds

Example: Weak and strong duality

A nonconvex problem with strong duality
Primal problem

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

$A \not\succeq 0$, hence nonconvex
Dual function

$$g(\lambda) = \inf_x \left(x^T (A + \lambda I) x + 2b^T x - \lambda \right) \tag{3}$$

1. unbounded below if $A + \lambda I \not\succeq 0$ or if $A + \lambda I \succeq 0$ and $b \notin \mathcal{R}(A + \lambda I)$
2. minimized by $x = -(A + \lambda I)^\dagger b$ otherwise: $g(\lambda) = -b^T (A + \lambda I)^\dagger b - \lambda$

Dual problem and equivalent SDP:

$$\begin{array}{ll} \texttt{minimize} & -b^T (A + \lambda I)^\dagger b - \lambda \\ \texttt{s.t.} & A + \lambda I \succeq 0 \\ & b \in \mathcal{R}(A + \lambda I) \end{array} \qquad \begin{array}{ll} \texttt{maximize} & -t - \lambda \\ \texttt{s.t.} & \begin{bmatrix} A + \lambda I & b \\ b^T & t \end{bmatrix} \succeq 0 \end{array}$$

Strong duality although primal problem is not convex (not easy to show)

4 Optimality Conditions

Complementary slackness: assume strong duality holds, x^* is primal optimal, (λ^*, ν^*) is dual optimal

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

hence, the two inequalities hold with equality

1. x^* minimizes $L(x, \lambda^*, \nu^*)$
2. $\lambda_i^* f_i(x^*) = 0$ for $i = 1, \ldots, m$ (known as complementary slackness):

$$\lambda_i^* > 0 \Longrightarrow f_i(x^*) = 0, \; f_i(x^*) < 0 \Longrightarrow \lambda_i^* = 0$$

4 Optimality Conditions

Karush-Kuhn-Tucker(KKT) conditions

The following four conditions are called KKT conditions (for a problem with differentiable f_i, h_i):

1. primal constraints: $f_i(x) \le 0, i = 1, \ldots, m$, $h_i(x) = 0$, $i = 1, \ldots, p$
2. dual constraints: $\lambda \succeq 0$
3. complementary slackness: $\lambda_i f_i(x) = 0, i = 1, \ldots, m$
4. gradient of Lagrangian with respect to x vanishes:

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

from page 5–17: if strong duality holds and x, λ, ν are optimal, then they must satisfy the KKT conditions

4 Optimality Conditions

(Sufficient) KKT conditions for convex problem

if $\tilde{x}, \tilde{\lambda}, \tilde{\mu}$ satisfy KKT for a convex problem, then they are optimal:

1. from complementary slackness and primal feasibility: $f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$

2. from 4th condition(and convexity): $g(\tilde{\lambda}, \tilde{\nu}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$

hence, $f_0(\tilde{x}) = g(\tilde{\lambda}, \tilde{\nu})$

(Moreover) if Slater' s condition is satisfied:

x is optimal if and only if there exist λ, ν that satisfy KKT conditions

1. recall that Slater implies strong duality, and dual optimum is attained

2. generalizes optimality condition $\nabla f_0(x) = 0$ for unconstrained problem

*5 Perturbation and sensitivity analysis

(unperturbed) optimization problem and its dual

$$\begin{array}{ll} \texttt{minimize} & f_0(x) \\ \texttt{subject to} & f_i(x) \leq 0, i = 1, \ldots, m \\ & h_i(x) = 0, i = 1, \ldots, p \end{array} \qquad \begin{array}{ll} \texttt{maximize} & g(\lambda, \nu) \\ \texttt{subject to} & \lambda \succeq 0 \end{array}$$

perturbed problem and its dual

$$\begin{array}{ll} \texttt{min.} & f_0(x) \\ \texttt{s.t.} & f_i(x) \leq u_i, i = 1, \ldots, m \\ & h_i(x) = v_i, i = 1, \ldots, p \end{array} \qquad \begin{array}{ll} \texttt{max.} & g(\lambda, \nu) - u^T\lambda - v^T\nu \\ \texttt{s.t.} & \lambda \succeq 0 \end{array}$$

1. x is primal variable; u, v are parameters
2. $p^*(u, v)$ is optimal value as a function of u, v
3. We are interested in information about $p^*(u, v)$ that we can obtain from the solution of the unperturbed problem and its dual

*5 Perturbation and sensitivity analysis

global sensitivity result

assume strong duality holds for unperturbed problem, and that λ^*, ν^* are dual optimal for unperturbed problem

apply weak duality to perturbed problem:

$$\begin{aligned} p^*(u, v) &\geq g(\lambda^*, \nu^*) - u^T\lambda^* - v^T\nu^* \\ &= p^*(0, 0) - u^T\lambda^* - v^T\nu^* \end{aligned}$$

sensitivity interpretation

1. if λ_i^* large: p^* increases greatly if we tighten constraint i ($u_i < 0$)
2. if λ_i^* small: p^* does not decrease much if we loosen constraint i ($u_i > 0$)
3. if ν_i large and positive: p^* increases greatly if we take $v_i < 0$; if ν_i large and negative: p^* increases greatly if we take $v_i > 0$
4. if ν_i small and positive: p^* does not decrease much if we take $v_i > 0$; if ν_i^* small and negative: p^* does not decrease much if we take $v_i < 0$

*5 Perturbation and sensitivity analysis

local sensitivity: if (in addition) $p^*(u, v)$ is differentiable at $(0, 0)$, then

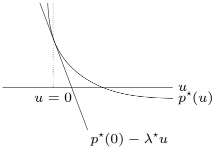
$$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i}, \qquad \nu_i^* = -\frac{\partial p^*(0,0)}{\partial v_i}$$

proof. (for λ_i^*): from global sensitivity result,

$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \searrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t} \geq -\lambda_i^*$$
$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \nearrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t} \leq -\lambda_i^*$$

hence, equality

$p^*(u)$ for a problem with one(inequality) constraint:



6 Examples

Problem reformulationsn and duality

1. equivalent formulations of a problem can lead to very different duals
2. reformulating the primal problem can be useful when the dual is difficult to derive, or uninteresting

Common reformulations

1. introduce new variables and equality constraints
2. make explicit constraints implicit or vice-versa
3. transform objective or constraint functions e.g., replace $f_0(x)$ by $\phi(f_0(x))$ with ϕ convex, increasing

6 Examples

Example 1. Introducing new variables and equality constraints

$$\texttt{minimize} \quad f_0(Ax+b)$$

1. dual function is constant: $g = \inf_x L(x) = \inf_x f_0(Ax+b) = p^*$
2. we have strong duality, but dual is quite useless

reformulated problem and its dual

$$\begin{array}{ll} \texttt{minimize} & f_0(y) \\ \texttt{subject to} & Ax+b-y=0 \end{array} \qquad \begin{array}{ll} \texttt{maximize} & b^T\nu - f_0^*(\nu) \\ \texttt{subject to} & A^T\nu = 0 \end{array}$$

dual function follows from

$$\begin{aligned} g(\nu) &= \inf_{x,y}(f_0(y) - \nu^T y + \nu^T Ax + b^T\nu) \\ &= \begin{cases} -f_0^*(\nu) + b^T\nu & A^T\nu = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

6 Examples

Example 2. norm approximation problem: minimize $\|Ax-b\|$

$$\begin{array}{ll} \texttt{minimize} & \|y\| \\ \texttt{subject to} & y = Ax - b \end{array}$$

can look up conjugate of $\|\cdot\|$, or derive dual directly

$$\begin{aligned} g(\nu) &= \inf_{x,y}(\|y\| - \nu^T y - \nu^T Ax + b^T \nu) \\ &= \begin{cases} -b^T\nu + \inf_y(\|y\| + \nu^T y) & A^T\nu = 0 \\ -\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} -b^T\nu & A^T\nu = 0, \|\nu\|_* \leq 1 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

dual of norm approximation problem

$$\begin{array}{ll} \texttt{maximize} & b^T\nu \\ \texttt{subject to} & A^T\nu = 0, \|\nu\|_* \leq 1 \end{array}$$

6 Examples

Example 3. Implicit constraints

LP with box constraints: primal and dual problem

$$\begin{array}{ll} \texttt{minimize} & c^T x \\ \texttt{subject to} & Ax + b = 0 \\ & -\mathbf{1} \preceq x \preceq \mathbf{1} \end{array} \qquad \begin{array}{ll} \texttt{maximize} & -b^T \nu - \mathbf{1}^T \lambda_1 - \mathbf{1}^T \lambda_2 \\ \texttt{subject to} & c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\ & \lambda_1 \succeq 0, \lambda_2 \succeq 0 \end{array}$$

reformulation with box constraints made implicit

$$\begin{array}{ll} \texttt{minimize} & f_0(x) = \begin{cases} c^T x & -\mathbf{1} \preceq x \preceq \mathbf{1} \\ \infty & \text{otherwise} \end{cases} \\ \texttt{subject to} & Ax = b \end{array}$$

dual function

$$\begin{aligned} g(\nu) &= \inf_{-\mathbf{1} \preceq x \preceq \mathbf{1}} (c^T x + \nu^T (Ax - b)) \\ &= -b^T \nu - \|A^T \nu + c\|_1 \end{aligned}$$

dual problem: maximize $-b^T \nu - \|A^T \nu + c\|_1$

6 Examples

Example 4. (Find duality) Figure out LP formulation and dual LP of the following Piecewise-linear minimization problem:

$$\texttt{minimize } f(x) = \max_{i=1,\dots,m} \left(a_i^T x + b_i\right)$$

Solution: LP problem is $\texttt{min.}\, t \ \texttt{s.t.}\ a_i^T x + b_i \leqslant t \quad i = 1, \cdots, m$, with Lagrangian:

$$L(t, x, \lambda) = t + \sum_{i=1}^{m} \lambda_i \left(a_i^T x + b_i - t\right) = t - \sum_{i=1}^{m} \lambda_i t + \left(\sum_{i=1}^{m} \lambda_i a_i^T\right) x + \sum_{i=1}^{m} b_i^T \lambda_i$$

$$= \left(1 - \sum_{i=1}^{m} \lambda_i\right) t + \left(\sum_{i=1}^{m} \lambda_i a_i^T\right) x + \sum_{i=1}^{m} b_i^T \lambda_i$$

6 Examples

Then

$$g(\lambda) = \inf_{t,x} \left(1 - \sum_{i=1}^{m} \lambda_i\right) t + \left(\sum_{i=1}^{m} \lambda_i a_i^T\right) x + \sum_{i=1}^{m} b_i^T \lambda_i$$

$$= \sum_{i=1}^{m} b_i^T \lambda_i + \inf_t \left(1 - \sum_{i=1}^{m} \lambda_i\right) t + \inf_x \left(\sum_{i=1}^{m} \lambda_i a_i^T\right) x$$

$$\text{if } \sum_{i=1}^{m} \lambda_i a_i^T \neq 0, g(\lambda) = -\infty; \text{ if } 1 - \sum_{i=1}^{m} \lambda_i \neq 0, g(\lambda) = -\infty$$

Then, dual LP is

$$\begin{array}{ll} \texttt{maximize} & \sum_{i=1}^{m} b_i^T \lambda_i \\ \texttt{subject to} & \sum_{i=1}^{m} \lambda_i a_i^T = 0 \\ & \sum_{i=1}^{m} \lambda_i = 1 \\ & \lambda_i \geq 0 \quad i = 1, \cdots, m \end{array}$$

6 Examples

Example 5. (Use dual formulation to solve primal problem) Maximize entropy:

$$\begin{array}{ll} \texttt{minimize} & f_0(x) = \sum_{i=1}^{n} x_i \log x_i \\ \texttt{subject to} & Ax \preceq b \\ & \mathbf{1}^T x = 1 \end{array}$$

6 Examples

Example 5. (Use dual formulation to solve primal problem) Maximize entropy:

$$\begin{array}{ll} \texttt{minimize} & f_0(x) = \sum_{i=1}^n x_i \log x_i \\ \texttt{subject to} & Ax \preceq b \\ & \mathbf{1}^T x = 1 \end{array}$$

Solution:

$$\begin{array}{ll} \texttt{maximize} & -b^T\lambda - \nu - e^{-\nu-1} \sum_{i=1}^n e^{-a_i^T \lambda} \\ \texttt{subject to} & \lambda \succeq 0 \end{array}$$

6 Examples

Example 6. (Use dual formulation to solve primal problem) Minimize summarized functions with equality constraints:

$$\begin{array}{ll} \texttt{minimize} & f_0(x) = \sum_{i=1}^n f_i(x_i) \\ \texttt{subject to} & a^T x = b \end{array}$$

f_i is strictly convex and differentiable

Solution: Lagrangian is

$$L(x, \nu) = \sum_{i=1}^n f_i(x_i) + \nu(a^T x - b) = -b\nu + \sum_{i=1}^n (f_i(x_i) + \nu a_i x_i)$$

$g(\nu) = -b\nu - \sum_{i=1}^n f_i^*(-\nu a_i)$ and the dual problem is

$$\texttt{maximize} \quad -b\nu - \sum_{i=1}^n f_i^*(-\nu a_i)$$

6 Examples

Example 7. (Use KKT conditions to solve QP)

$$\begin{array}{ll} \texttt{minimize} & 1/2x^TPx + q^Tx + r \\ \texttt{subject to} & Ax = b \end{array}$$

Solution: write down the KKT conditions

$$Ax^* = b, \qquad Px^* + q + A^T\nu^* = 0$$

6 Examples

Example 7. (Use KKT to solve Bandwidth Maximization prob.)

$$\begin{array}{ll} \texttt{minimize} & -\sum_{i=1}^{n} \log(\alpha_i + x_i) \\ \texttt{subject to} & x \succeq 0, \ \mathbf{1}^T x = 1 \end{array}$$

Solution: write down the KKT conditions

$$\begin{aligned} x^* \succeq 0, \quad \mathbf{1}^T x^* = 1, \quad \lambda^* \succeq 0, \quad \lambda_i^* x_i^* = 0, \ i = 1, \dots, n \\ -1/(\alpha_i + x_i^*) - \lambda_i^* + \nu^* = 0, \ i = 1, \dots, n \end{aligned}$$

Rewrite KKT conditions:

$$\begin{aligned} x^* \succeq 0, \quad \mathbf{1}^T x^* = 1, \quad x_i^*(\nu^* - 1/(\alpha_i + x_i^*)) = 0, \ i = 1, \dots, n \\ \nu^* \geq 1/(\alpha_i + x_i^*), \ i = 1, \dots, n \end{aligned}$$

6 Examples

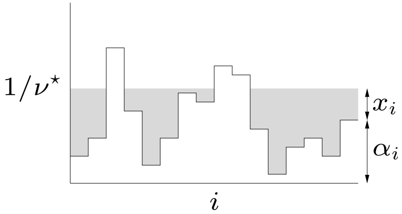
(1) if $\nu^* < 1/\alpha_i$, $x_i^* > 0$, and $x_i^* = 1/\nu^* - \alpha_i$;

(2) if $\nu^* \geq 1/\alpha_i$, $x_i^* \leq 0$, i.e., $x_i^* = 0$

As a result, $x_i^* = \max\{0, 1/\nu^* - \alpha_i\}$

Determine the optimal value of ν combined with $\mathbf{1}^T x^* = 1$, that is

$$\sum_{i=1}^{n} \max\{0, 1/\nu^* - \alpha_i\} = 1$$



6 Examples

Example 8. Use KKT conditions to find the point in the following set which is the closest to $(0,0)$.

$$M = \left\{x \in \mathbb{R}^2 : x_1 + x_2 \geq 4, 2x_1 + x_2 \geq 5\right\}$$

Solution: make it an optimization problem:

$$\begin{aligned} \texttt{minimize} \quad & x_1^2 + x_2^2 \\ \texttt{subject to} \quad & -x_1 - x_2 + 4 \leq 0, \\ & -2x_1 - x_2 + 5 \leq 0. \end{aligned}$$

The Lagrangian is

$$\begin{aligned} & L\left(x_1, x_2, u_1, u_2\right) \\ =& x_1^2 + x_2^2 + u_1\left(-x_1 - x_2 + 4\right) + u_2\left(-2x_1 - x_2 + 5\right), u_1, u_2 \geq 0. \end{aligned}$$

6 Examples

The KKT conditions are:

i) Feasibility

ii) Complementary Slackness

$$
\begin{aligned}
u_1 \left(-x_1 - x_2 + 4\right) = 0, u_1 \geq 0 \\
u_2 \left(-2x_1 - x_2 + 5\right) = 0, u_2 \geq 0
\end{aligned}
$$

iii) Optimality

$$
\begin{aligned}
\frac{\partial L}{\partial x_1} =& 2x_1 - u_1 - 2u_2 = 0, \\
\frac{\partial L}{\partial x_2} =& 2x_2 - u_1 - u_2 = 0
\end{aligned}
$$

6 Examples

case (1): $u_1 = 0, u_2 = 0$. From iii), $x_1 = 0, x_2 = 0$, not feasible;
case (2): $x_1 + x_2 = 4, u_2 = 0$. Combined with iii), we have

$$2x_1 - u_1 = 0,$$
$$2x_2 - u_1 = 0,$$
$$x_1 + x_2 = 4,$$

which yield $x_1 = 2, x_2 = 2, u_1 = 4 > 0$. A KKT point $(2, 2, 4, 0)$!
case (3): $u_1 = 0,\ 2x_1 + x_2 = 5$.

$$2x_1 - 2u_2 = 0,$$
$$2x_2 - u_2 = 0,$$
$$2x_1 + x_2 = 5$$

We have $x_1 = 2, x_2 = 1, u_2 = 2$ (not feasible).
case (4): $x_1 + x_2 = 4, 2x_1 + x_2 = 5$. we have $x_1 = 1, x_2 = 3$ and

$$u_1 + 2u_2 = 2,$$
$$u_1 + u_2 = 6$$

which yield $u_1 = 10, u_2 = -4 < 0$ (not satisfy dual feasibility).

6 Examples

Example 9. Following is nonconvex optimization problem with strong duality holding。How to find the optimal solution, unique?

$$\begin{array}{ll} \texttt{maximize}_x & -2(x_1-2)^2 - x_2^2 \\ \texttt{subject to} & x_1^2 + x_2^2 \leq 25 \\ & x_1 \geq 0 \end{array}$$

Solution: the Lagrangian is

$$\mathcal{L}\left(x_1, x_2, \mu_1, \mu_2\right) = -2\left(x_1-2\right)^2 - x_2^2 + \mu_1\left(x_1^2 + x_2^2 - 25\right) - \mu_2 x_1$$

KKT conditions:

$$\begin{cases} -4\left(x_1-2\right) + 2\mu_1 x_1 - \mu_2 = 0 \\ -2x_2 + 2\mu_1 x_2 = 0 \\ x_1^2 + x_2^2 \leq 25 \\ x_1 \geq 0 \\ \mu_1, \mu_2 \geq 0 \\ \mu_1 = 0 \quad \text{or} \quad x_1^2 + x_2^2 = 25 \\ \mu_2 = 0 \quad \text{or} \quad x_1 = 0 \end{cases}$$

6 Examples

(1) $\mu_1 = \mu_2 = 0$. We have $x_1 = 2$ and $x_2 = 0$, satisfying primal constraints. Hence, $(2, 0)$ and $(0, 0)$ are possible solutions with the value equal to 0.

(2) $\mu_1 = 0$ and $\mu_2 > 0$. We have $x_1 = x_2 = 0$ and $\mu_2 = 8 > 0$. The value is -8.

(3) $\mu_2 = 0$ and $\mu_1 > 0$, we have

$$\begin{cases} -4(x_1 - 2) + 2\mu_1 x_1 = 0 \\ -2x_2 + 2\mu_1 x_2 = 0 \\ x_1 \geq 0 \\ \mu_1 > 0 \\ x_1^2 + x_2^2 = 25 \end{cases}$$

If $x_2 = 0$, $x_1 = 5, x_2 = 0$, and $\mu_1 = 6/5 > 0$ and $\mu_2 = 0$, get a value of -18. If $x_2 \neq 0$, $\mu_1 = 1$. we have $x_1 = 4$ and $x_2 = \pm 3$, and $\mu_2 = 0$, we get a value of -17.

(4) $\mu_2 > 0$ and $\mu_1 > 0$. We have $x_1 = 0$ and $x_2 = \pm 5$, $\mu_1 = 1$ and $\mu_2 = 8$, get a value of -33, which is the minimum.