Primal-Dual Theory

Zhe Wang
Applied Mathematics
西安交通大學
XI'AN JIAOTONG UNIVERSITY

2018 年 9 月 27 日

Max-min characterization of weak and strong duality

primal problem and dual problem

$$\begin{aligned} &\text{minimize} \quad f_0(x) \\ &\text{subject to } f_i(x) \leq 0, \quad i = 1, \cdots, m \end{aligned} \tag{1}$$

it is equivalent to the problem

$$\min_x \max_{\lambda_i \geq 0} \; f_0(x) + \lambda_i f_i(x) \tag{2}$$

dual problem:

$$\max_{\lambda_i \geq 0} \min_x \; f_0(x) + \lambda_i f_i(x) \tag{3}$$

the inequality

$$p* = \min_x \max_{\lambda_i \geq 0} \; f_0(x) + \lambda_i f_i(x) \geq \max_{\lambda_i \geq 0} \min_x \; f_0(x) + \lambda_i f_i(x) = d^*$$

always comes ture.

when equality holds, we claim f satisfies saddle point property

saddle point interpretation

Definition (saddle point)

We call a pair of points $(\hat{w}, \hat{z})$ is the saddle point of f, if the inequality

$$f(\hat{w}, z) \leq f(\hat{w}, \hat{z}) \leq f(w, \hat{z}) \tag{4}$$

holds for all $w \in W,\ z \in Z$, which means that $\hat{w}$ minimize $f(w, \hat{z})$ and $\hat{z}$ maximize $f(w, \hat{z})$.

In other words, the strong min-max property holds, and the common value is $f(\hat{w}, \hat{z})$. Returning to our discussion of Lagrange duality, we see that if $\hat{x}$ and $\hat{\lambda}$ are primal and dual optimal points for a problem in which strong duality obtains, they form a saddle-point for the Lagrangian, the converse is also true if $\hat{\lambda} \geq 0$.

game interptretation

we give an interesting interpretation of such max-min or saddle point problem in terms of a continuous zero-sum problem:

rule of the game

two players: player 1, player 2;

player 1 choose the first variable $w \in W$, while player 2 select the second one $z \in Z$;

player 1 pays an amount $f(w, z)$ to player 2;

player 1 wants to minimize $f(w, z)$ as player 2 wants to maximize f.

based on the rule, there are two subproblems which comes from who will be the first player:

condition 1:

suppose player 1 makes his choice fist, then comes the player 2. player 2 wants to maximize the payoff $f(w, z)$, and so will choose $z \in Z$ to maximize $f(w, z)$. The resulting payoff will be $\sup_{z \in Z} f(w, z)$, which depends on w, the choice of the first player.

Player 1 knows (or assumes) that player 2 will follow this strategy, and so will choose $w \in W$ to make this worst-case payoff to player 2 as small as possible. Thus player 1 chooses

$$\arg\min_{w \in W} \sup_{z \in Z} f(w, z) \tag{5}$$

which result in the payoff

$$\inf_{w \in W} \sup_{z \in Z} f(w, z) \tag{6}$$

condition 2

now suppose the order of the play is reversed, which means Player 2 must choose $z \in Z$ first, and then player 1 chooses $w \in W$, following a similar argument, the pay off from player 1 to player 2 will be

$$\sup_{z \in Z} \inf_{w \in W} f(w, z) \tag{7}$$

from the max-min inequality, it is always better for a player to go second, in which case, the player has a prior knowledge of his opponent's strategy.

if the payoff function $f(w, z)$ is the lagrangian, domain $W = \mathbb{R}^n, Z = \mathbb{R}^m_+$, player 1 makes his choice first , the minimum payoff for player 1 will be p^*, if the player 2 comes first, the minimum payoff for player 1 will be d^*, that is to say, the optimal duality gap for the problem is exactly equal to the advantage afforded the player who goes second.

certificate of suboptimality and stopping criteria

main conclusion

Recall: for any point (λ, ν) that is dual feasible, the lower bound of the $g(\lambda, \nu)$can be given by the optimal value of the primal problem:

$$p^* \geq g(\lambda, \nu) \tag{8}$$

the strong duality means the existence of the (λ, ν) such that the equality holds.

now comes to our conclusion: if x is primal feasible, we have

$$f_0(x) - p* \leq f_0(x) - g(\lambda, \nu) \tag{9}$$

this establishes that x is ϵ -suboptimal, with $\epsilon = f_0(x) - g(\lambda, \nu)$

suppose (λ, ν) is dual feasible, x is primal feasible,

$$p* \in [g(\lambda, \nu), f_0(x)], \;\; d^* \in [g(\lambda, \nu), f_0(x)] \tag{10}$$

the existence of the (λ, ν) and x such that the strong duality holds can be think of as a certificate that x is optimal and (λ, ν) is dual optimal.

for optimization algorithm, suppose we produce a sequence of primal feasible $\{x_k\}$ and dual feasible (λ_k, ν_k) and $\epsilon_{abs} > 0$ is a given required accuracy, then the stopping criteria :

$$f_0(x_k) - g(\lambda_k, \nu_k) \leq \epsilon_{abs} \tag{11}$$

guarantee x_k is ϵ_{abs} suboptimal.

complementary slackness

suppose x^* is primal optimal (λ^*, ν^*) is dual optimal and the strong duality holds, this means:

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

conclusion

$$
\sum_{i=1}^{m} \lambda_i^* f_i(x^*) = 0 \tag{12}
$$

it is not difficult to get $\lambda_i^* f_i(x^*) = 0$, this condition is known as complementary slackness.

KKT optimality conditions

Assume all functions $f_0, \cdots, f_m, h_1, \cdots, h_p$ are differentiable, still no assumption of convexity.

Karush-Kuhn-Tucker conditions

let $x^*, (\lambda^*, \nu^*)$ be primal optimal and dual optimal, respectively, the duality gap of these two points are 0. Now, the following conditions hold:

- $f_i(x^*) \leq 0, i = 1, \cdots, m$
- $h_i(x^*) = 0, i = 1, \cdots, p$
- $\lambda_i^* \geq 0, i = 1, \cdots, m$
- $\lambda_i^* f_i(x^*) = 0, i = 1, \cdots, m$
- $\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x^*) = 0$

to summarize, any optimization problem with differentiable objective and constraint functions for which the strong duality obtains, any pair of primal and dual optimal points satisfy KKT conditions.

KKT conditions for convex problems

Main conclusion

If $f_i(x)$ are all convex and $h_i(x)$ are affine, $\hat{x}, \hat{\lambda}, \hat{\nu}$ are points satisfy the KKT conditions, then $\hat{x}$,$(\hat{\lambda}, \hat{\nu})$ are optimal for two problems, respectively, with zero duality gap.

proof

$$\begin{aligned} g(\hat{\lambda}, \hat{\nu}) &= L(\hat{x}, \hat{\lambda}, \hat{\nu}) \\ &= f_0(\hat{x}) + \sum_{i=1}^{m} \hat{\lambda}_i f_i(\hat{x}) + \sum_{i=1}^{p} \hat{\nu}_i h_i(\hat{x}) \\ &= f_0(\hat{x}) \end{aligned} \tag{13}$$

based on the discussion, if a convex optimization problem with differentiable objective and constraint functions satisfies Slater's condition, then the KKT conditions provide necessary and sufficient conditions for optimality

Example

$$\begin{aligned} \text{minimize} \quad & -\sum_{i=1}^{n} \log(\alpha_i + x_i) \\ \text{subject to} \quad & x_i \geq 0, \sum x_i = 1; \end{aligned} \tag{14}$$

Solution

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

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

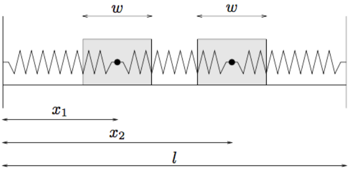
if $\nu^* < \dfrac{1}{\alpha_i}$, *then* $x_i^* > 0$, *which means* $\nu^* = 1/(\alpha_i + x_i^*)$, *the assumption* $\nu^* \geq \dfrac{1}{\alpha_i}$ *would violates the complementary slackness unless* $x_i^* = 0$

$$x_i^\star = \begin{cases} 1/\nu^\star - \alpha_i & \nu^\star < 1/\alpha_i \\ 0 & \nu^\star \geq 1/\alpha_i, \end{cases}$$

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

the lefthand side is a piecewise-linear increasing function, with breakpoints at α_i, thus the unique solution exists.

Mechanics interpretation of KKT conditions



$$\begin{array}{ll} \text{minimize} & (1/2)\left(k_1x_1^2 + k_2(x_2-x_1)^2 + k_3(l-x_2)^2\right) \\ \text{subject to} & w/2 - x_1 \le 0 \\ & w + x_1 - x_2 \le 0 \\ & w/2 - l + x_2 \le 0, \end{array}$$

$$\lambda_1(w/2 - x_1) = 0, \qquad \lambda_2(w - x_2 + x_1) = 0, \qquad \lambda_3(w/2 - l + x_2) = 0,$$

$$\begin{bmatrix} k_1x_1 - k_2(x_2-x_1) \\ k_2(x_2-x_1) - k_3(l-x_2) \end{bmatrix} + \lambda_1\begin{bmatrix} -1 \\ 0 \end{bmatrix} + \lambda_2\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \lambda_3\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0.$$

the same equations can be get from Newton's three basic principles .

Solving the primal problem via the dual

If the strong duality holds, and (λ^*, ν^*) is the dual optimal, then any primal optimal point is also a minimizer of $L(x, \lambda^*, \nu^*)$

Now, suppose the strong duality holds, the (λ^*, ν^*) is dual optimal, considering the following problem:

$$\arg\text{minimize} \quad f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x) \tag{15}$$

in some condition, the solution is unique (if it is a strictly convex function of x).

If the solution is primal feasible, it is the primal optimal; if not, we claim the primal optimal can't be attained.

Example (separable function subject to an equality constraint)

$$
\begin{aligned}
&\text{minimize} \quad f_0(x) = \sum_{i=1}^{n} f_i(x_i) \\
&\text{subject to} \quad a^{\mathrm{T}} x = b
\end{aligned}
\tag{16}
$$

Solution

The Lagrangian is

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

which is also separable, so the dual function is

$$
\begin{aligned}
g(\nu) &= -b\nu + \inf_x \left(\sum_{i=1}^{n} (f_i(x_i) + \nu a_i x_i) \right) \\
&= -b\nu + \sum_{i=1}^{n} \inf_{x_i} (f_i(x_i) + \nu a_i x_i) \\
&= -b\nu - \sum_{i=1}^{n} f_i^*(-\nu a_i).
\end{aligned}
$$

The dual problem is thus

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

with (scalar) variable $\nu \in \mathbf{R}$.

Now suppose we have found an optimal dual variable $\nu^\star$. (There are several simple methods for solving a convex problem with one scalar variable, such as the bisection method.) Since each f_i is strictly convex, the function $L(x, \nu^\star)$ is strictly convex in x, and so has a unique minimizer $\tilde{x}$. But we also know that $x^\star$ minimizes $L(x, \nu^\star)$, so we must have $\tilde{x} = x^\star$. We can recover $x^\star$ from $\nabla_x L(x, \nu^\star) = 0$, *i.e.*, by solving the equations $f_i'(x_i^\star) = -\nu^\star a_i$.

the perturbed problem

Definition (the perturbed problem)

we claim the following problem is the perturbed version of the primal problem:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq u_i, i = 1, \cdots, m \\ & h_i(x) = v_i, i = 1, \cdots, p. \end{aligned} \tag{17}$$

such perturbed of the primal problem can be seen as the relaxed($i > 0$) or the tightened($u_i < 0$) version of ith inequality for the primal problem.

we define the $p^*(u, v) : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ as the optimal value of the (u,v-perpured problem), then have $p^*(0, 0) = p^*$

a global inequality

with the assumption that the strong duality holds, the dual optimum is attained (for example, the original problem is convex, and the Slater's condition is satisfied), (λ^*, ν^*) be the optimal for the unperturbed problem, we have:

$$p^*(u, v) \geq p^*(0, 0) - \lambda^{*\mathrm{T}} u - \nu^{*\mathrm{T}} v \tag{18}$$

proof:

$$\begin{aligned} p*(0,0) = g(\lambda^*, \nu^*) &\leq f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{\nu_i}^{*} h_i(x) \\ &\leq f_0(x) + \lambda^{*\mathrm{T}} u + \nu^{*\mathrm{T}} v \end{aligned} \tag{19}$$

sensitivity interpretations

Based on the inequality, we can analyze the sensitivity of the optimal Lagrange variables, with the assumption strong duality holds. Some conclusions are listed as follows:

- if λ_i^* is large, with $i < 0$, the optimal value will increase greatly;
- if nu_i^* is large and positive, with $v_i < 0$, the optimal value will increase greatly, the results for the reverse can be discussed analogously.
- if λ_i^* is small, the loosen of the ith constraint won' t result in too much decrease of the $p^*(u, v)$
- if nu_i^* is small and positive, with $v_i > 0$, then the optimal value $p^*(u, v)$ will not decrease too much.

Local sensitivity analysis

main conclusion

Suppose $p^*(u,v)$ is differentiable at $u=0, v=0$, then with the strong duality holds, we have:

$$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i}, \quad \nu_i^* = -\frac{\partial p^*(0,0)}{\partial \nu_i} \tag{20}$$

now, you can give the local sensitivity analysis of the $p(u,v)$ at a small neighbor of $(0,0)$

tightening the ith inequality constraint a small amount (i.e. taking u_i small and negative) yields an increase in p^* of approximately $-\lambda_i^* u^*$, while loosing the ith constraint a small amount yields a decrease in p^* of approximately $\lambda_i^* u_i$

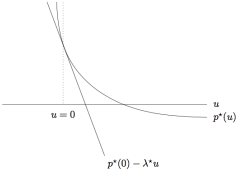


Figure 5.10 Optimal value $p^\star(u)$ of a convex problem with one constraint $f_1(x) \le u$, as a function of u. For $u=0$, we have the original unperturbed problem; for $u<0$ the constraint is tightened, and for $u>0$ the constraint is loosened. The affine function $p^\star(0)-\lambda^\star u$ is a lower bound on $p^\star$.

local sensitivity analysis

Proof.

$$\lim_{t \to 0} \frac{p^*(te_i, 0) - p^*}{t} = \frac{\partial p^*(0,0)}{\partial u_i} \tag{21}$$

while the inequality indicates that for $t > 0$

$$\frac{p^*(te_i, 0) - p^*}{t} \geq -\lambda_i^*, \text{which indicates} \tag{22}$$

$$\frac{\partial p^*(0,0)}{\partial u_i} \geq -\lambda_i^* \tag{23}$$

while taking the limit $t < 0$ yields the opposite inequality, so

$$\frac{\partial p^*(0,0)}{\partial u_i} = -\lambda_i^* \tag{24}$$

□

The local sensitivity result (5.58) gives us a quantitative measure of how active a constraint is at the optimum x.

If $f_i(x) < 0$, then the constraint is inactive,the small disturbance won't affect the optimal value. (complementary slackness)

But now suppose that $f_i(x) = 0$, i.e., the ith constraint is active at the optimum. The ith optimal Lagrange multiplier tells us how active the constraint is.

If λ_i is small, it means that the constraint can be loosened or tightened a bit without much effect on the optimal value.

Introducing new variables and equality constraints

Sometimes, for the convenience of the calculate, it is necessary to introducing the new variables, while the cost is the introducing of the equality constrains.

Examples

consider the problem:

$$\text{minimize} \quad f_0(Ax+b) \tag{25}$$

It's Lagrange dual function is a constant p^*, which indicates the strong duality holds, i.e., $p^* = d^*$, but is non of use. the reformulation of the problem by introducing the new variables is:

$$\text{minimize} \quad f_0(y) \tag{26}$$

$$\text{subject to} \quad Ax+b=y \tag{27}$$

the Lagrangian of the reformulated problem is

$$L(x,y,\nu) = f_0(x) + \nu(Ax+b-y) \tag{28}$$

minimize $L(x,y,\nu)$ over x and y

Examples

$$g(\nu) = b^{\mathrm{T}}\nu + \inf_y (f_0(y) - \nu^{\mathrm{T}} y) = b^{\mathrm{T}}\nu - f_0^*(\nu) \tag{29}$$

where $f_0^*(x)$ is the conjugate of f_0. the dual problem can therefor be expressed as

$$\text{maximize} \quad b^{\mathrm{T}}\nu - f_0^*(\nu) \tag{30}$$

$$\text{subject to} \quad A^{\mathrm{T}}\nu = 0 \tag{31}$$

it is considerably more useful than the dual of the original problem

Examples

norm approximation problem

$$\text{minimize} \quad ||Ax - b|| \tag{32}$$

we reformulate the problem as :

$$\text{minimize} ||y|| \tag{33}$$

$$\text{subject to} Ax - b = y \tag{34}$$

the Lagrangian is

$$\text{maximize} \quad b^{\mathrm{T}}\nu \tag{35}$$

$$\text{subject to} \quad ||\nu||_* \leq 1 \tag{36}$$

$$A^{\mathrm{T}}\nu = 0 \tag{37}$$

the idea of introducing the new variables can be applied to the constraint functions as well considering the problem:

$$\text{minimize} \quad f_0(Ax_0 + b_0) \tag{38}$$

$$\text{subject to} \quad f_i(A_i x + b_i) \leq 0, i = 1, \cdots, m \tag{39}$$

we can reformulate the problem as:

$$\text{minimize} \quad f_0(y_0) \tag{40}$$

$$\text{subject to} \quad f_i(y_i) \leq 0, i = 1, \cdots, m \tag{41}$$

$$A_i x + b_i = y_i \tag{42}$$

following the same steps, we will get the dual problem of the reformulated problem is

$$\text{maximize} \quad \sum_{i=0}^{m} \nu_i^{\mathrm{T}} b_i - f_0^*(\nu_0) - \sum_{i=1}^{m} \lambda_i f_i^*(\nu_i/\lambda_i) \tag{43}$$

$$\text{subject to} \quad \lambda_i > 0 \tag{44}$$

$$\sum_{i=0}^{m} A_i^{\mathrm{T}} \nu_i = 0 \tag{45}$$

Transforming the objectives

Sometimes we need to transform the objective f_0, for example, to an increasing function of f_0 (the equivalence of the two problems can be seen in 4.1.3), if the dual problems of the new one can provide more information:

Examples (minimum norm problem)

$$\text{minimize} \quad ||Ax - b|| \tag{46}$$

Reformulation of the problem:

$$\text{minimize} \quad 1/2||y||^2 \tag{47}$$

$$\text{subject to} \quad Ax - b = y \tag{48}$$

Dual problem of the new one:

$$\text{maximize} \quad -1/2||\nu||_*2 + b^T\nu \tag{49}$$

$$\text{subject to} \quad A^T\nu = 0 \tag{50}$$

Implicit constraints

the main idea is that we can introduce some of the constraints in the objective function, by modifying the objective function to be infinite when the constraint is violated.

Examples

linear program with box constriants

$$\begin{aligned} \text{minimize} \quad & c^{\mathrm{T}}x && (51) \\ \text{subject to} \quad & Ax = b && (52) \\ & l_i \leq x_i \leq u_i \text{sometimes called box constraints or variable bounds} && (53) \end{aligned}$$

the dual problem:

$$\begin{aligned} \text{maximize} \quad & -b^{\mathrm{T}}\nu - \lambda_1^{\mathrm{T}}u + \lambda_2^{\mathrm{T}}l && (54) \\ \text{subject to} \quad & A^{\mathrm{T}}\nu + \lambda_1 - \lambda_2 + c = 0 && (55) \\ & \lambda_1^i \geq 0, \lambda_2^i \geq 0 && (56) \end{aligned}$$

reformulation problems:

$$\text{minimize} \quad f_0(x) \tag{57}$$

$$\text{subject to} \quad Ax = b \tag{58}$$

where we define:

$$f_0(x) = \begin{cases} c^{\mathrm{T}}x & l \leq x \leq u \\ \infty & \text{otherwise} \end{cases} \tag{59}$$

the dual function:

$$g(\nu) = \inf_{l \leq x \leq u} (c^{\mathrm{T}}x + \nu^{\mathrm{T}}(Ax - b) \tag{60}$$

$$= -b^{\mathrm{T}}\nu - u^{\mathrm{T}}(A^{\mathrm{T}}\nu + c)^{-} + l^{\mathrm{T}}(A^{\mathrm{T}}\nu + c)^{+} \tag{61}$$

the dual problem:

$$\text{maximize} \quad -b^{\mathrm{T}}\nu - u^{\mathrm{T}}(A^{\mathrm{T}}\nu + c)^{-} + l^{\mathrm{T}}(A^{\mathrm{T}}\nu + c)^{+} \tag{62}$$

Weak alternatives via the dual function

In this subsection, we focus on the problem of determining feasibility of a system

$$\begin{aligned} f_i(x) &\leq 0, \quad i = 1, \cdots, m \\ h_i(x) &= 0, \quad i = 1, \cdots, p \end{aligned} \tag{63}$$

the problem:

$$\begin{aligned} \text{minimize} \quad & 0 \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \\ & h_i(x) = 0, \quad i = 1, \cdots, p \end{aligned} \tag{64}$$

has the optimal value

$$p^* = \begin{cases} 0, & \text{the problem(64) is feasible} \\ \infty, & \text{the problem(64) is infeasible} \end{cases} \tag{65}$$

so solving the optimization problem (64) is the same as solving the inequality system (63).

the dual problem

dual problem of (64)

$$g(\lambda,\nu)=\inf_{x\in D}\Big(\sum_{i=1}^{m}\lambda_i f_i(x)+\sum_{i=1}^{p}\nu_i h_i(x)\Big) \tag{66}$$

based on the homogeneous of $g(\lambda,\nu)$,

$$d^*=\begin{cases}\infty, & \lambda\geq 0, g(\lambda,\nu)>0 \quad \text{is feasible}\\ 0, & \lambda\geq 0, g(\lambda,\nu)>0 \quad \text{is infeasible}\end{cases} \tag{67}$$

Weak duality indicates that $d^*\leq p^*$, so at most one of two inequality system:

$$system1 \qquad \begin{aligned} f_i(x)\leq 0, &\quad i=1,\cdots,m\\ h_i(x)=0, &\quad i=1,\cdots,p\end{aligned} \qquad system2 \qquad \lambda\geq 0, \quad g(\lambda,\nu)>0 \tag{68}$$

is feasible. The condition is called weak alternatives

Strict inequalities

the feasibility of the strict inequality:

$$f_i(x) < 0, \quad h_i(x) = 0 \tag{69}$$

the alternative inequality system is

$$\lambda \geq 0, \quad \lambda \neq 0, \quad g(\lambda, \nu) \geq 0 \tag{70}$$

strict inequalities

Strong alternatives, which means exactly one of the two alternative inequality systems holds, can be get if $f_i(x)$ are convex and h_i are affine,and some type of constraint qualification holds.
in this subsection, we assume $f_i(x)$ are convex, while $h_i(x)$ are affine. strict inequality system:

$$f_i(x) < 0, Ax = b \tag{71}$$

alternatives:

$$\lambda \geq 0, \lambda \neq 0, g(\lambda, \nu) \geq 0 \tag{72}$$

Under the condition:

$$\text{exits } x \in relintD, \text{s.t. } Ax = b \tag{73}$$

the two inequality systems are strong alternatives.

the reformation of the primal system:

$$\text{minimize} \quad s \tag{74}$$

$$\text{subject to} \quad f_i(x) - s \leq 0, \quad i = 1, \cdots, m \tag{75}$$

$$Ax = b \tag{76}$$

The optimal value p^* of this problem is negative if and only if there is a solution to the strict inequality system

The Lagrange dual function for the problem is

$$\inf_{x \in D, s} \left(s + \sum_{i=1}^{m} \lambda_i (f_i(x) - s) + \nu^T (Ax - b)\right) = \begin{cases} g(\lambda, \nu) & 1^T \lambda = 1 \\ \infty & \text{otherwise} \end{cases} \tag{77}$$

therefore, the dual problem is :

$$\text{maximize} \quad g(\lambda, \nu) \tag{78}$$

$$\text{subject to} \quad \lambda \geq 0, 1^T \lambda = 1 \tag{79}$$

slater's condition holds for the problem(74), then we have $d^* = p^*$, in other words:

$$g(\lambda^*, \nu^*) = p^*, \quad \lambda^* \geq 0, \quad 1^T\lambda = 1 \tag{80}$$

If problem(71) is infeasible, $p^* \geq 0$, $g(\lambda^*, \nu^*) < 0$, which satisfies the inequalities in problem(72), then the problem(72) is feasible.

If the alternate system is feasible, $d^* = p^* \geq 0$, then the strict inequality system is infeasible

Intersection of ellipsoids

Main problem: when the intersection of m ellipsoids are not empty?

$$\epsilon_i = \{x | f_i(x) \leq 0\} \tag{81}$$

$$\text{with} \quad f_i(x) = x^{\mathrm{T}} A_i x + 2b_i^{\mathrm{T}} x + c_i, \; i = 1, \cdots, m \tag{82}$$

$$A_i \in S_{++}^n \tag{83}$$

this problem is equivalent to feasibility of strict quadratic inequalities:

$$f_i(x) = x^{\mathrm{T}} A_i x + 2b_i^{\mathrm{T}} x \tag{84}$$

Recall

Definition (proper cone)

A cone $K \subset \mathbb{R}^n$ is called a proper cone if:

- K is convex
- K is closed
- K is solid(nonempty interior)
- K is pointed(contains no line)

Definition (partial ordering)

$$x \leq_K y \Longleftrightarrow y - x \in K \tag{85}$$

$$x <_K y \Longleftrightarrow y - x \in \mathrm{int}K \tag{86}$$

all our problems and conclusions are generalized to the problem:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \preceq_{K_i} 0, \quad i = 1, \cdots, m \\ & h_i(x) = 0, \quad i = 1, \cdots, m \end{aligned} \tag{87}$$

with the constraint

$$\lambda_i \succeq_{K_i^*} 0, \quad i = 1, \cdots, m \tag{88}$$

K_i^* is the dual cone of K_i the weak duality holds the Lagrange dual problem is:

$$\text{maximize} \quad g(\lambda, \nu) \tag{89}$$

$$\text{subject to} \quad \lambda_i \succeq_{K_i^*} 0, i = 1, \cdots, m \tag{90}$$

Slater's condition and strong duality

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \preceq_{K_i} 0, \quad i = 1, \cdots, m \\ & Ax = b, \quad i = 1, \cdots, m \end{aligned} \tag{91}$$

if f_0 is convex, f_i are K_i-convex, there exists an $x \in \text{relint}\, D$ with $AX = b$, and $f_i(x) \prec_{K_i} 0$, this condition implies strong duality and the dual optimal can be attained

The conclusions of complementary slackness and KKT conditions can be extended to generalized problems as well:

complementary slackness

if the strong duality holds at (x^*, λ^*, ν^*)

$$\lambda_i^{*T} f_i(x^*) = 0 \tag{92}$$

from which, we can conclude:

$$\lambda_i^* >_{K_i^*} 0 \longrightarrow f_i(x^*) = 0, \quad f_i(x^*) <_{K_i^*} 0 \longrightarrow \lambda_i^* = 0 \tag{93}$$

KKT conditions

if all functions are differentiable, the problem satisfied KKT conditions:

- $$f_i(x^*) \preceq_{K_i} 0$$
- $$h_i(x^*) = 0$$
- $$\lambda_i^* \succeq_{K_i^*} 0$$
- $$\lambda_i^* f_i(x^*) = 0$$
- $$\nabla f_0(x^*) + \sum_{i=1}^{m} Df_i(x^*)^{\mathrm{T}} \lambda_i^* + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x^*) = 0$$

Perturbation and sensitivity analysis

the perturbed version of the generalized problem:

$$\text{minimize} \quad f_0(x) \tag{94}$$

$$\text{subject to} \quad f_i(x) \preceq_{K_i} u_i \tag{95}$$

$$h_i(x) = i \tag{96}$$

if the strong duality holds at x^*, λ^*, ν^*, then for all u and v:

$$p^*(u,v) \geq p^* - \sum_{i=1}^{m} \lambda_i^{*\mathrm{T}} u_i - \nu^{*\mathrm{T}} v \tag{97}$$

if $p^*(u,v)$ is differentiable at $u=0, v=0$, then

$$\lambda_i^* = -\nabla_{u_i} p^*(0,0) \tag{98}$$

Theorems of alternatives

generalized inequalities and equalities system

$$f_i(x) \preceq_{K_i} 0, \quad i = 1, \cdots, m \quad h_i(x) = 0, \quad i = 1, \cdots, p \tag{99}$$

strict inequalities system:

$$f_i(x) \prec_{K_i} 0, \quad i = 1, \cdots, m \quad h_i(x) = 0, \quad i = 1, \cdots, p \tag{100}$$

weak alternatives for the inequalities system:

$$\lambda_i \succeq_{K_i^*} 0, i = 1, \cdots, m, \quad g(\lambda, \nu) > 0 \tag{101}$$

the alternatives for the strictly inequalities system:

$$\lambda_i \succeq_{K_i^*} 0, i = 1, \cdots, m, \quad \lambda \neq 0 \quad g(\lambda, \nu) \geq 0 \tag{102}$$

strong alternatives

if f_i are K_i-convex, the functions h_i are affine, exits an $\hat{x} \in \text{relint}D, A\hat{x} = b$ strict inequality system:

$$f_i(x) <_{K_i} 0, \quad i = 1, \cdots, m \quad Ax = b \tag{103}$$

and its alternatives:

$$\lambda_i \geq_{K_i^*} 0, i = 1, \cdots, m, \quad \lambda \neq 0 \quad g(\lambda, \nu) \geq 0 \tag{104}$$

are strong alternatives