Chapter5_对偶

1. 拉格朗日对偶问题

原问题的拉格朗日函数
将约束条件加权后与原函数求和

Lagrangian: standard form problem (not necessarily convex)

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

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 $f_i(x)\le 0$

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

拉格朗日对偶函数
在x上对拉格朗日函数求下界。
g是凹函数，因为对偶函数是一族关于(λ，μ)的仿射函数的逐点下确界，即使原问题不是凸的，对偶函数也是凹函数

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

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

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

下界性质
对偶函数给出了原问题最优值p*的下界，即对任意(λ，μ)，g(λ，μ)都小于等于p*

求一个问题的对偶问题的一般方法
1. 写出原问题的拉格朗日函数
2. 对于原问题的优化变量，求拉格朗日函数的下确界（通常可以在梯度等于0处求得，也有用共轭函数的方法）
3. 将求得的x代回拉格朗日函数，得到拉格朗日对偶函数

例示

线性方程组的最小二乘问题
图中给出了线性方程组的最小二乘问题的对偶函数，以及下界性质在这个场景中的具体表现。

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

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

$$\nabla_xL(x,\nu)=2x+A^T\nu=0\implies x=-(1/2)A^T\nu$$

Plug into L to obtain g:

$$g(\nu)=L((-1/2)A^T\nu,\nu)=-\frac14\nu^TAA^T\nu-b^T\nu$$

A concave function of ν

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

标准形式的线性规划
对于线性函数，只有斜率为0时，下界才存在

Standard form LP

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

1. Lagrangian is

$$L(x,\lambda,\nu)=c^Tx+\nu^T(Ax-b)-\lambda^Tx=-b^T\nu+(c+A^T\nu-\lambda)^Tx$$

2. L is affine in x, hence

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

g is linear on affine domain $\{(\lambda,\nu)\mid A^T\nu-\lambda+c=0\}$, hence concave

Lower bound property: $p^*\ge-b^T\nu$ if $A^T\nu+c\succeq 0$

等式条件约束下的范数最小化
Equality constrained norm minimization

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

dual function:

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

where $\|v\|_*=\sup_{\|u\|\le1}u^Tv$ is dual norm of $\|\cdot\|$

proof: follows from $\inf_x(\|x\|-y^Tx)=0$ if $\|y\|_*\le 1$, $-\infty$ otherwise
1. If $\|y\|_*\le1$, $\inf_x(\|x\|-y^Tx)=\inf_{x\neq0}(1-y^Tx/\|x\|)\|x\|\ge0$, with equality if $x=0$
2. If $\|y\|_*>1$, choose $x=tu$ where $\|u\|\le1$, $u^Ty=\|y\|_*>1$:

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

Lower bound property: $p^*\ge b^T\nu$ if $\|A^T\nu\|_*\le1$

补充：对偶范数的定义
u是任意一个向量，对偶范数取遍所有小于等于1的向量。

这里的 $\|v\|_*$ 是 $\|\cdot\|$ 的对偶范数，定义为 $\|v\|_*=\sup_{\|u\|\le1}u^Tv$。

双向划分问题
Two-way partitioning

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

1. A non-convex problem; feasible set contains 2^n discrete points
2. interpretation: partition $\{1,\dots,n\}$ 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:

$$g(\nu)=\inf_x\left(x^TWx+\sum_i\nu_i(x_i^2-1)\right)=\inf_x x^T(W+\mathbf{diag}(\nu))x-\mathbf{1}^T\nu=\begin{cases}-\mathbf{1}^T\nu & W+\mathbf{diag}(\nu)\succeq0\\ -\infty & \text{otherwise}\end{cases}$$

Lower bound property: $p^*\ge-\mathbf{1}^T\nu$ if $W+\mathbf{diag}(\nu)\succeq0$

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

拉格朗日对偶和共轭函数的关系
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:

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

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

拉格朗日对偶问题：找出原问题最优值的最大下界
The dual problem

$$\begin{array}{ll}\texttt{maximize} & g(\lambda,\nu)\\ \texttt{subject to} & \lambda\succeq0\end{array}$$

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 $\lambda\succeq0$, $(\lambda,\nu)\in\mathbf{dom}\,g$
4. Often simplified by making implicit constraint $(\lambda,\nu)\in\mathbf{dom}\,g$ explicit

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

$$\begin{array}{ll}\texttt{minimize} & c^Tx\\ \texttt{subject to} & Ax=b\\ & x\succeq0\end{array}\qquad\begin{array}{ll}\texttt{maximize} & -b^T\nu\\ \texttt{subject to} & A^T\nu+c\succeq0\end{array}$$

2. 弱对偶与强对偶

弱对偶与强对偶
弱对偶性一定成立，强对偶性则通常不成立，但是强对偶性对于凸问题**通常**成立。能够保证强对偶性成立的条件称为**约束准则**。

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

$$\begin{array}{ll}\texttt{maximize} & -\mathbf{1}^T\nu\\ \texttt{subject to} & W+\mathbf{diag}(\nu)\succeq0\end{array}$$

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**

Slater约束准则（充分条件，即Slater条件成立推出强对偶性成立，但是强对偶性成立不能推出Slater条件成立）
只对**凸问题**有效；
在定义域的非空内部存在一点，使得所有约束条件**严格**成立，则Slater约束准则成立，进而保证强对偶性成立；
Slater约束准则还保证了最优值有下界；
还有很多其他的约束准则。

Slater’s constraint qualification

Strong duality holds for a convex problem

$$\begin{array}{ll}\texttt{maximize} & -f_0(x)\\ \texttt{subject to} & f_i(x)\le0,\ i=1,\dots,m\\ & Ax=b\end{array}$$

if it is strictly feasible, i.e.,

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

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

2.there exist many other types of constraint qualifications

Slater条件成立的情况下强对偶性成立的证明
Assumptions: 1. contain non-empty interior point; 2. **rank** $A=p$. Define the following two exclusive sets:

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

$\exists(\tilde\lambda,\tilde\nu,\mu)\neq0$ and α, such that

$$(u,v,t)\in\mathcal{A}\implies\tilde\lambda^Tu+\tilde\nu^Tv+\mu t\ge\alpha$$
$$(u,v,t)\in\mathcal{B}\implies\tilde\lambda^Tu+\tilde\nu^Tv+\mu t\le\alpha$$

We have 1. $\tilde\lambda\succeq0$ and $\mu\ge0$ (due to feature of epigraph); 2. $\mu p^*\le\alpha$

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

$$\sum\tilde\lambda_if_i(x)+\tilde\nu^T(Ax-b)+\mu f_0(x)\ge\alpha\ge\mu p^*$$

续
Case 1: $\mu>0$

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

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

Case 2: $\mu=0$

$$\sum\tilde\lambda_if_i(x)+\tilde\nu^T(Ax-b)\ge0,\forall x\in\mathcal{D}$$

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

$$\sum\tilde\lambda_if_i(\tilde x)\ge0\implies\tilde\lambda=0$$

As a result, $\tilde\nu\neq0$ (non-zero parameters for a hyperplane). Consider the following facts: 1. $\tilde\nu^T(A\tilde x-b)=0$, 2. $\tilde x\in\mathbf{int}\,\mathcal{D}$. There exists x, such that $\tilde\nu^T(Ax-b)<0$ unless $A^T\tilde v=0$. (**rank**$A=p$, $\tilde\nu$ is p-dimension).

续：几何意义
上述证明的几何意义如图5.6所示，图中的问题是具有一个不等式约束的简单问题。分离集合 $\mathcal{A}$ 和集合 $\mathcal{B}$ 的超平面在点 $(0,p^*)$ 处定义了 $\mathcal{A}$ 的一个支撑超平面。根据 Slater 约束准则，分离的超平面必须是非竖直的（即存在一个形如 $(\lambda^*,1)$ 的法向量）。（针对具有一个不等式约束的凸优化问题，强对偶性也有可能不成立，习题 5.21 给出了一个简单的例子。）

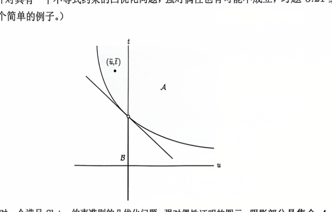


图5.6 对一个满足 Slater 约束准则的凸优化问题，强对偶性证明的图示。阴影部分是集合 $\mathcal{A}$，加粗的竖直线段是集合 $\mathcal{B}$，不包含点 $(0,p^*)$，图中空心圆点所示。两个集合都是凸集，不相交，因此存在分离超平面。根据 Slater 约束准则，任意分离超平面必是非竖直的，这是因为它必须穿过点 $(\tilde u,\tilde t)=(f_1(\tilde x),f_0(\tilde x))$ 的左侧，其中点 $\tilde x$ 严格可行。

例示

带不等式约束的线性规划问题
Inequality form LP

Primal problem

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

Dual function

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

Dual problem

$$\begin{array}{ll}\texttt{maximize} & -b^T\lambda\\ \texttt{subject to} & A^T\lambda+c=0,\ \lambda\succeq0\end{array}$$

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

二次规划问题
该问题对偶函数的求得，符合对X求导后将导数置为0再代回拉格朗日函数的一般范式。

Quadratic program

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

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

Dual function

$$g(\lambda)=\inf_x\left(x^TPx+\lambda^T(AX-B)\right)=-\frac14\lambda^TAP^{-1}A^T\lambda-b^T\lambda\quad(2)$$

Dual problem

$$\begin{array}{ll}\texttt{maximize} & -(1/4)\lambda^TAP^{-1}A^T\lambda-b^T\lambda\\ \texttt{subject to} & \lambda\succeq0\end{array}$$

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

非凸但满足强对偶性的问题的例子
A nonconvex problem with strong duality

Primal problem

$$\begin{array}{ll}\texttt{minimize} & x^TAx+2b^Tx\\ \texttt{subject to} & x^Tx\le1\end{array}$$

$A\not\succeq0$, hence nonconvex

Dual function

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

1. unbounded below if $A+\lambda I\not\succeq0$ or if $A+\lambda I\succeq0$ 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:

$$\begin{array}{ll}\texttt{minimize} & -b^T(A+\lambda I)^\dagger b-\lambda\\ \texttt{s.t.} & A+\lambda I\succeq0\\ & b\in\mathcal{R}(A+\lambda I)\end{array}\qquad\begin{array}{ll}\texttt{maximize} & -t-\lambda\\ \texttt{s.t.} & \begin{bmatrix}A+\lambda I & b\\ b^T & t\end{bmatrix}\succeq0\end{array}$$

maximize

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

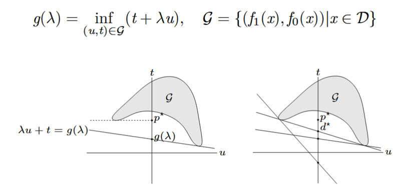
3. 几何解释

简单起见，考虑只有一个不等式约束的问题
p*：原问题的最优解。由于原问题f1(x)<=0，因此只能在u<=0上做优化。
d*：对偶问题的最优解。对偶问题不考虑f1的正负，因此在x的整个定义域中求上界。

For simplicity, consider problem with one constraint $f_1(x)\le0$

interpretation of dual function:

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

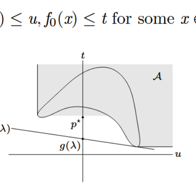


1. $\lambda u+t=g(\lambda)$ is (non-vertical) supporting hyperplane to $\mathcal{D}$
2. hyperplane intersects t-axis at $t=g(\lambda)$

上镜图版本
针对具有一个（不等式）约束的问题，对偶函数和下界的几何解释。

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

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



4. 最优条件

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

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

hence, the two inequalities hold with equality
1. x^* minimizes $L(x,\lambda^*,\nu^*)$
2. $\lambda_i^*f_i(x^*)=0$ for $i=1,\dots,m$ (known as complementary slackness):

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

通过以上等式得到两个结论：
1. 拉格朗日函数在x*处取得最小值；
2. x* 处有图中所示的结论。这个结论就叫做**互补松弛性**，该式意味着除了第i个约束起作用的情况下，最优Lagrange乘子的第i项都为0。

KKT最优性条件
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: $f_i(x)\le0,i=1,\dots,m$, $h_i(x)=0$, $i=1,\dots,p$
2. dual constraints: $\lambda\succeq0$
3. complementary slackness: $\lambda_if_i(x)=0,i=1,\dots,m$
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

KKT最优性条件：
1. 原问题的约束成立；
2. 对偶问题的约束成立；
3. 互补松弛性成立；
4. 原函数导数为0.
满足以上4个条件的点称为**KKT最优点**，如果强对偶性成立且x、λ、μ都是最优的，那么它一定满足KKT条件。
反过来，我们可以通过求出所有的KKT最优点处的函数值，值最小的就是最优点。

对于凸问题，如果一个点满足了KKT条件，那么它就是最佳的。

如果Slater约束条件满足，当且仅当存在满足KKT 条件的 λ、ν 时，X 是最优的。

对于无约束问题，KKT条件泛化为导数为0.

*5. 扰动及灵敏度分析

6. 例示

本节通过一些例子说明，对一个问题进行简单的等价变形可能得到很不一样的对偶问题。

考虑如下几种类型的变形：
1. 引入新的变量以及相应的等式约束。
2. 用原目标函数的增函数取代原目标函数。
3. 将显式的约束隐式表达，例如将其并入目标函数的定义域。

当原问题的对偶问题难以推导时，重写原问题是非常有用的。

例1.引入新变量以及相应的等式约束

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

1. dual function is constant: $g=\inf_xL(x)=\inf_xf_0(Ax+b)=p^*$
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

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

例2.范数逼近问题
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

$$\begin{aligned}g(\nu)&=\inf_{x,y}(\|y\|-\nu^Ty-\nu^TAx+b^T\nu)\\&=\begin{cases}-b^T\nu+\inf_y(\|y\|+\nu^Ty) & A^T\nu=0\\ -\infty & \text{otherwise}\end{cases}\\&=\begin{cases}-b^T\nu & A^T\nu=0,\|\nu\|_*\le1\\ -\infty & \text{otherwise}\end{cases}\end{aligned}$$

dual of norm approximation problem

$$\begin{array}{ll}\texttt{maximize} & b^T\nu\\ \texttt{subject to} & A^T\nu=0,\|\nu\|_*\le1\end{array}$$

例3.隐式约束
Example 3. Implicit constraints

LP with box constraints: primal and dual problem

$$\begin{array}{ll}\texttt{minimize} & c^Tx\\ \texttt{subject to} & Ax+b=0\\ & -\mathbf{1}\preceq x\preceq\mathbf{1}\end{array}\qquad\begin{array}{ll}\texttt{maximize} & -b^T\nu-\mathbf{1}^T\lambda_1-\mathbf{1}^T\lambda_2\\ \texttt{subject to} & c+A^T\nu+\lambda_1-\lambda_2=0\\ & \lambda_1\succeq0,\lambda_2\succeq0\end{array}$$

reformulation with box constraints made implicit

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

dual function

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

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

例4.无穷范数最小化的LP形式以及对偶LP
Example 4. (Find duality) Figure out LP formulation and dual LP of the following Piecewise-linear minimization problem:

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

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

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

续
Then

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

If $\sum_{i=1}^m\lambda_ia_i^T\neq0$, $g(\lambda)=-\infty$; If $1-\sum_{i=1}^m\lambda_i\neq0$, $g(\lambda)=-\infty$

Then, dual LP is

$$\begin{array}{ll}\texttt{maximize} & \sum_{i=1}^mb_i^T\lambda_i\\ \texttt{subject to} & \sum_{i=1}^m\lambda_ia_i^T=0\\ & \sum_{i=1}^m\lambda_i=1\\ & \lambda_i\ge0\quad i=1,\dots,m\end{array}$$

例5.（用对偶问题解决原问题）熵的最大化
Example 5. (Use dual formulation to solve primal problem) Maximize entropy:

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

Solution:

$$\begin{array}{ll}\texttt{maximize} & -b^T\lambda-\nu-e^{-\nu-1}\sum_{i=1}^ne^{-a_i^T\lambda}\\ \texttt{subject to} & \lambda\succeq0\end{array}$$

例6.（用对偶问题解决原问题）最小化等式约束下的求和问题
Example 6. (Use dual formulation to solve primal problem) Minimize summarized functions with equality constraints:

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

f_i is strictly convex and differentiable

Solution: Lagrangian is

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

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

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

例7.用KKT条件解决问题

例7.1 解决QP问题
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

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

例7.2 解决带宽最大化问题
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

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

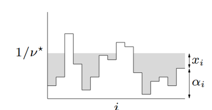
续
(1) if $\nu^*<1/\alpha_i$, $x_i^*>0$, and $x_i^*=1/\nu^*-\alpha_i$

(2) if $\nu^*\ge1/\alpha_i$, $x_i^*\le0$, i.e., $x_i^*=0$

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

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

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



例8.用KKT条件找出集合中最接近原点的点
Example 8. Use KKT conditions to find the point in the following set which is the closest to (0,0).

$$M=\{x\in\mathbb{R}^2: x_1+x_2\ge4,\ 2x_1+x_2\ge5\}$$

Solution: make it an optimization problem:

$$\begin{array}{ll}\texttt{minimize} & x_1^2+x_2^2\\ \texttt{subject to} & -x_1-x_2+4\le0,\\ & -2x_1-x_2+5\le0.\end{array}$$

The Lagrangian is

$$L(x_1,x_2,u_1,u_2)=x_1^2+x_2^2+u_1(-x_1-x_2+4)+u_2(-2x_1-x_2+5),\ u_1,u_2\ge0$$

续
The KKT conditions are:

i) Feasibility

ii) Complementary Slackness

$$u_1(-x_1-x_2+4)=0,\ u_1\ge0$$
$$u_2(-2x_1-x_2+5)=0,\ u_2\ge0$$

iii) Optimality

$$\frac{\partial L}{\partial x_1}=2x_1-u_1-2u_2=0,\qquad\frac{\partial L}{\partial x_2}=2x_2-u_1-u_2=0$$

续
case (1): $u_1=0,u_2=0$. From (ii), $x_1=0,x_2=0$, not feasible;

case (2): $x_1+x_2=4$, $u_2=0$. Combined with iii), we have

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

which yield $x_1=2,x_2=2,u_1=4>0$. A KKT point $(2,2,4,0)$!

case (3): $u_1=0$, $2x_1+x_2=5$.

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

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

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

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

例9.用KKT条件找出非凸但强对偶问题的最优解
Example 9. Following is nonconvex optimization problem with strong duality holding. How to find the optimal solution, unique?

$$\begin{array}{ll}\texttt{maximize}_x & -2(x_1-2)^2-x_2^2\\ \texttt{subject to} & x_1^2+x_2^2\le25\\ & x_1\ge0\end{array}$$

Solution: the Lagrangian is

$$\mathcal{L}(x_1,x_2,\mu_1,\mu_2)=-2(x_1-2)^2-x_2^2+\mu_1(x_1^2+x_2^2-25)-\mu_2x_1$$

KKT conditions:

$$\begin{cases}-4(x_1-2)+2\mu_1x_1-\mu_2=0\\-2x_2+2\mu_1x_2=0\\x_1^2+x_2^2\le25\\x_1\ge0\\\mu_1,\mu_2\ge0\\\mu_1=0\ \text{ or }\ x_1^2+x_2^2=25\\\mu_2=0\ \text{ or }\ x_1=0\end{cases}$$

续
(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

$$\begin{cases}-4(x_1-2)+2\mu_1x_1=0\\-2x_2+2\mu_1x_2=0\\x_1\ge0\\\mu_1>0\\x_1^2+x_2^2=25\end{cases}$$

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 $x_2\neq0$, $\mu_1=1$. we have $x_1=4$ and $x_2=\pm3$, and $\mu_2=0$, we get a value of -17.

(4) $\mu_2>0$ and $\mu_1>0$. We have $x_1=0$ and $x_2=\pm5$, $\mu_1=1$ and $\mu_2=8$, get a value of -33, which is the minimum.