中国農業大學 China Agricultural University

对偶问题

翟卫欣 副教授

zhaiweixin@cau.edu.cn

中国农业大学信息与电气工程学院

目录

目录

约束优化的对偶问题

为什么要研究对偶问题?

线性规划的对偶问题举例1

原问题：

某工厂生产A_1, A_2两类产品，它们都需要在B_1, B_2, B_3三类不同的设备上加工，所需时间如下，如何安排生产计划可以使该厂利润最大？

设备	单位台时		总有限台时
	A_1	A_2	
B_1	3	4	36
B_2	5	4	40
B_3	9	8	76
利润	32	30	

单位台时：单位产品(A_1, A_2)的加工台时

线性规划的对偶问题举例1

$$
\begin{aligned}
max\ z &= 32x_1 + 30x_2 \\
s.t.\ 3x_1 + 4x_2 &\leqslant 36 \\
5x_1 + 4x_2 &\leqslant 40 \\
9x_1 + 8x_2 &\leqslant 76 \\
x_1, x_2 &\geqslant 0
\end{aligned}
$$

线性规划的对偶问题举例1

另一位老板要租用该工厂的机器，设y_1, y_2, y_3为租用B_1, B_2, B_3三种设备的机器的费用，该老板希望最小化租用成本w，但是该经营者所花费的租用费用不能太低，至少应该不低于生产产品A_1, A_2的利润。

$$
\begin{aligned}
min\ w &= 36y_1 + 40y_2 + 76y_3 \\
s.t.\ 3y_1 + 5y_2 + 9y_3 &\geqslant 32 \\
4y_1 + 4y_2 + 8y_3 &\geqslant 30 \\
y_1, y_2, y_3 &\geqslant 0
\end{aligned}
$$

线性规划的对偶问题举例2

原问题：

$$
\begin{aligned}
max\ & 2x_1 + 3x_2 \\
s.t.\ & 4x_1 + 8x_2 \leqslant 12 \\
& 2x_1 + 1x_2 \leqslant 3 \\
& 3x_1 + 2x_2 \leqslant 4 \\
& x_1, x_2 \geqslant 0
\end{aligned}
$$

为了找到原问题的上限，我们可以根据约束条件构造原问题

$$2x_1 + 3x_2 \leqslant 4x_1 + 8x_2 \leqslant 12$$

$$2x_1 + 3x_2 \leqslant \frac{1}{2}(4x_1 + 8x_2) \leqslant 6$$

$$2x_1 + 3x_2 \leqslant \frac{1}{3}((4x_1 + 8x_2) + (2x_1 + 1x_2) \leqslant 5$$

线性规划的对偶问题举例2

三个不等式约束条件都乘以$y_i \geqslant 0, i = 1, 2, 3$
使得三式分别乘以y_i累加后对应的x_1的系数大于等于2 使得三式分别乘以y_i累加后对应的x_3的系数大于等于3 此时所对应的$12y_1 + 3y_2 + 4y_3$必然比原问题的最优解要大，所以尽量最小化$12y_1 + 3y_2 + 4y_3$即可

$$
\begin{aligned}
min\ & 12y_1 + 3y_2 + 4y_3 \\
s.t.\ & 4y_1 + 2y_2 + 3y_3 \geqslant 2 \\
& 8y_1 + 1y_2 + 2y_3 \geqslant 3 \\
& y_1, y_2, y_3 \geqslant 0
\end{aligned}
$$

当$x_1 = \frac{1}{2}, x_2 = \frac{5}{4}$时或者$y_1 = \frac{5}{16}, y_2 = 0, y_3 = \frac{1}{4}$时，原问题和对偶问题都得到最小值4.75。

线性规划的对偶问题

对称形式的线性规划问题特征：
(1)全部约束条件为不等式
(1.1)对极大化问题的约束条件都是$\leqslant$
(1.2)对极小化问题的约束条件都是$\geqslant$
(2)全部变量为非负
原问题$\Rightarrow$对偶问题

$$
\begin{aligned}
max\ z &= c_1x_1 + c_2x_2 + \cdots + c_nx_n \\
s.t.\ a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &\leqslant b_1 \\
a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &\leqslant b_2 \\
&\vdots \\
a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &\leqslant b_m \\
x_1, x_2, \cdots, x_n &\geqslant 0
\end{aligned}
\Rightarrow
\begin{aligned}
min\ w &= b_1y_1 + b_2y_2 + \cdots + b_my_m \\
s.t.\ a_{1,1}y_1 + a_{2,1}y_2 + \cdots + a_{m,1}y_m &\geqslant c_1 \\
a_{1,2}y_1 + a_{2,2}y_2 + \cdots + a_{m,2}y_m &\geqslant c_2 \\
&\vdots \\
a_{1,n}y_1 + a_{2,n}y_2 + \cdots + a_{m,n}y_m &\geqslant c_n \\
y_1, y_2, \cdots y_m &\geqslant 0
\end{aligned}
$$

线性规划的对偶问题

	x_1	x_2	$\cdots$	x_n	原始约束	$min\ w$
y_1	$a_{1,1}$	$a_{1,2}$	$\cdots$	$x_{1,n}$	$\leqslant$	b_1
y_2	$a_{2,1}$	$a_{2,2}$	$\cdots$	$x_{2,n}$	$\leqslant$	b_2
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\leqslant$	$\cdots$
y_m	$a_{m,1}$	$a_{m,2}$	$\cdots$	$x_{m,n}$	$\leqslant$	b_m
对偶约束	$\geqslant$	$\geqslant$	$\cdots$	$\geqslant$		
$max\ z$	c_1	c_2	$\cdots$	c_n		

原问题 对偶问题

$$
\begin{aligned}
max\ z &= \boldsymbol{c}^T\boldsymbol{x} & \qquad min\ w &= \boldsymbol{b}^T\boldsymbol{y} \\
s.t. \boldsymbol{A}\boldsymbol{x} &\leqslant \boldsymbol{b} \ \ \Rightarrow & s.t. \boldsymbol{A}^T\boldsymbol{y} &\geqslant \boldsymbol{c} \\
\boldsymbol{x} &\geqslant \boldsymbol{0} & \boldsymbol{y} &\geqslant \boldsymbol{0}
\end{aligned}
$$

目录

约束优化的对偶问题

考虑一般形式的约束优化问题：

$$
\begin{aligned}
(P)\ min\, f(\boldsymbol{x}),& \\
s.t.\ g_i(\boldsymbol{x}) &\geqslant 0, i=1,\cdots,m; \\
h_i(\boldsymbol{x}) &= 0, i=1,\cdots,l; \\
\boldsymbol{x} &\in \boldsymbol{D}.
\end{aligned}
$$

- P即*primal problem*，原问题；
- $\boldsymbol{D}$是集合约束，如$\boldsymbol{D}=\mathbb{R}^n$，$\boldsymbol{X}=\mathbb{Z}^n$(整数规划)，$\boldsymbol{X}=\{0,1\}$（0-1规划）。如果将问题写成只有等式约束和不等式约束的情况，则集合约束默认为$\boldsymbol{D}=\mathbb{R}^n$。

约束优化的对偶问题

对于对偶问题的研究，常需要引入拉格朗日函数，定义目标函数为：

$$\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) = inf\{f(\boldsymbol{x}) - \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) - \sum_{i=1}^{l} \mu_i h_i(\boldsymbol{x}) | \boldsymbol{x} \in \boldsymbol{D}\}$$

inf:下界，sup:上界

原问题的对偶问题为

$$max\ \theta(\boldsymbol{\lambda}, \boldsymbol{\mu})$$

$$s.t.\ \boldsymbol{\lambda} \geqslant \boldsymbol{0}$$

对偶函数为$\theta(\boldsymbol{\lambda}, \boldsymbol{\mu})$，部分情况下$\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) = -\infty$

约束优化的对偶问题

原问题和对偶问题的矩阵形式

记$\boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), \cdots, g_m(\boldsymbol{x}))^T$， $\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \cdots, h_l(\boldsymbol{x}))^T$

则原问题为：

$$
\begin{aligned}
&min\, f(\boldsymbol{x}),\\
&s.t.\ \boldsymbol{g}(\boldsymbol{x}) \geqslant \boldsymbol{0}\\
&\quad\ \ \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}\\
&\quad\ \ \boldsymbol{x} \in \boldsymbol{D}.
\end{aligned}
$$

对偶问题为：

$$
\begin{aligned}
&max\ \theta(\boldsymbol{\lambda}, \boldsymbol{\mu}),\\
&s.t.\ \boldsymbol{\lambda} \geqslant \boldsymbol{0}
\end{aligned}
$$

其中对偶函数为

$$\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) = inf\{f(\boldsymbol{x}) - \boldsymbol{\lambda}^T\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\mu}^T\boldsymbol{h}(\boldsymbol{x}) | \boldsymbol{x} \in \boldsymbol{D}\}$$

约束优化的对偶问题

考虑非线性规划问题：

$$
\begin{aligned}
\min \quad & x_1^2 + x_2^2 \\
\text{s.t.} \quad & x_1 + x_2 - 4 \geq 0, \\
& x_1, x_2 \geq 0.
\end{aligned}
$$

解:将变量的非负限制作为集约束，即

$$
\boldsymbol{x} \in D = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \middle| x_1 \geq 0, x_2 \geq 0 \right\},
$$

对偶函数为

$$
\begin{aligned}
\theta(w) &= \inf \left\{ x_1^2 + x_2^2 - w(x_1 + x_2 - 4) | x_1 \geq 0, x_2 \geq 0 \right\} \\
&= \inf \left\{ x_1^2 - w x_1 | x_1 \geq 0 \right\} + \inf \left\{ x_2^2 - w x_2 | x_2 \geq 0 \right\} + 4w.
\end{aligned}
$$

由上式可知，当$w \geq 0$时，有

$$\theta(w) = -\frac{1}{2}w^2 + 4w.$$

当$w < 0$时，由于$x_1 \geq 0, x_2 \geq 0$,则有

$$x_1^2 - wx_1 \geq 0,$$
$$x_2^2 - wx_2 \geq 0.$$

因此，当$x_1 = x_2 = 0$时，得到极小值

$$\theta(w) = 4w.$$

综上分析，得到对偶函数

$$\theta(w) = \begin{cases} -\dfrac{1}{2}w^2 + 4w, w \geq 0, \\ 4w, w < 0. \end{cases}$$

本例的对偶问题为

$$\begin{aligned} \max \quad & -\frac{1}{2}w^2 + 4w \\ \text{s.t.} \quad & w \geq 0. \end{aligned}$$

不难求得原问题的最优解

$$\bar{\boldsymbol{x}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix},$$

目标函数的最优值$f_{min} = 8$，而对偶问题的最优解$\overline{w} = 4$，最优值$\theta_{max} = 8$.

弱对偶定理

弱对偶定理：设$\boldsymbol{x}$和$(\boldsymbol{\lambda},\boldsymbol{\mu})$分别是原问题和对偶问题的可行解，则
$f(\boldsymbol{x}) \geqslant \theta(\boldsymbol{\lambda},\boldsymbol{\mu})$
证明：根据$\theta(\boldsymbol{\lambda},\boldsymbol{\mu})$的定义，有
$\theta(\boldsymbol{\lambda},\boldsymbol{\mu}) = inf\{f(\boldsymbol{y}) - \boldsymbol{\lambda}^T\boldsymbol{g}(\boldsymbol{y}) - \boldsymbol{\mu}^T\boldsymbol{h}(\boldsymbol{y})|\boldsymbol{y} \in \boldsymbol{D}\} \leqslant f(\boldsymbol{x}) - \boldsymbol{\lambda}^T\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\mu}^T\boldsymbol{h}(\boldsymbol{x})$
由于$\boldsymbol{x}$和$(\boldsymbol{\lambda},\boldsymbol{\mu})$分别是原问题和对偶问题的可行解，即满足
$\boldsymbol{g}(\boldsymbol{x}) \geqslant \boldsymbol{0}, \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$和$\boldsymbol{\lambda} \geqslant \boldsymbol{0}$，
所以得$f(\boldsymbol{x}) \geqslant \theta(\boldsymbol{\lambda},\boldsymbol{\mu})$

弱对偶定理的推论

推论1：对于原问题和对偶问题，必有
$inf\{f(\boldsymbol{x})|\boldsymbol{g}(\boldsymbol{x}) \geqslant \boldsymbol{0}, \boldsymbol{h}(\boldsymbol{x}) = 0, \boldsymbol{x} \in \boldsymbol{D}\} \geqslant sup\{\theta(\boldsymbol{\lambda}, \boldsymbol{\mu})|\boldsymbol{\lambda} \geqslant \boldsymbol{0}\}$
推论2：如果$f(\bar{\boldsymbol{x}}) \leqslant \theta(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}})$，其中$\bar{\boldsymbol{x}} \in \{\boldsymbol{x}|\boldsymbol{g}(\boldsymbol{x}) \geqslant \boldsymbol{0}, \boldsymbol{h}(\boldsymbol{x}) = 0, \boldsymbol{x} \in \boldsymbol{D}\}$，$\bar{\boldsymbol{\lambda}} \geqslant \boldsymbol{0}$，则$\bar{x}$和$(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}})$分别是原问题和对偶问题的最优解。
推论3：如果$inf\{f(\boldsymbol{x})|\boldsymbol{g}(\boldsymbol{x}) \geqslant \boldsymbol{0}, \boldsymbol{h}(\boldsymbol{x}) = 0, \boldsymbol{x} \in \boldsymbol{D}\} = -\infty$，则对每一个$\boldsymbol{\lambda} \geqslant \boldsymbol{0}$，有$\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) = -\infty$
推论4：如果$sup\{\theta(\boldsymbol{\lambda}, \boldsymbol{\mu})|\boldsymbol{\lambda} \geqslant \boldsymbol{0}\} = \infty$，则原问题没有可行解

目录

对偶间隙

由弱对偶定理的推论1可知原问题的目标函数的最优值f_{min}和对偶问题的目标函数的最优值θ_{max}满足关系$f_{min} \geqslant \theta_{max}$，如果严格不等号成立，则称存在对偶间隙(Duality Gap)。

$Duality\ Gap = f_{min} - \theta_{max}$

对偶间隙举例

求解如下问题的对偶间隙。

$$\begin{aligned} min \quad & x_1^2 + x_2^2 \\ s.t. \quad & -x_1 - x_2 \leqslant -\frac{1}{2} \\ & x_1, x_2 \in \mathbb{Z}_+ \end{aligned}$$

其中$\mathbb{Z}_+$代表$0, 1, 2, \cdots$

原问题$f_{min} = 1$，在$\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \ or \begin{bmatrix} 0 \\ 1 \end{bmatrix}$时取到。

对偶间隙举例

对偶问题

$$\theta(\boldsymbol{\lambda}) = \min_{\boldsymbol{x}\in\mathbb{Z}_+^2} \{x_1^2 + x_2^2 - \lambda(x_1 + x_2 - \frac{1}{2})\}$$

$$= \min_{\boldsymbol{x}\in\mathbb{Z}_+^2} \{(x_1 - \frac{\lambda}{2})^2 + (x_2 - \frac{\lambda}{2})^2 + \frac{\lambda}{2} - \frac{\lambda^2}{2}\}$$

$$= \begin{cases} \lambda/2 & \textit{if } 0 \leqslant \lambda \leqslant 1 \\ 2 - \frac{3}{2}\lambda & \textit{if } 1 < \lambda \leqslant 3 \\ 8 - \frac{7}{2}\lambda & \textit{if } 3 < \lambda \leqslant 5 \\ \dots \end{cases}$$

$$\theta_{max} = \frac{1}{2}, \ \lambda = 1$$

$$\textit{duality gap} \ = f_{min} - \theta_{max} = 1 - \frac{1}{2} = \frac{1}{2}$$

目录

中国农业大学 China Agricultural University

强对偶定理的引理

设$\boldsymbol{D}$是$\mathbb{R}^n$中的一个非空凸集，$\phi(\boldsymbol{x})$和$g_i(\boldsymbol{x})(i=1,2,\cdots,m)$分别是$\mathbb{R}^n$上的凸函数和凹函数，$h_j(\boldsymbol{x})(j=1,2,\cdots,l)$是$\mathbb{R}$上的线性函数，即假设

$$\boldsymbol{h}(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}$$

那么下列两个系统中，若系统1无解，则系统2有解$(\omega_0,\boldsymbol{\lambda},\boldsymbol{\mu})$；反之，若系统2有解$(\omega_0,\boldsymbol{\lambda},\boldsymbol{\mu})$且$\omega_0>0$，则系统1无解。

1. 系统1：存在$\boldsymbol{x}\in\boldsymbol{D}$，使得$\phi(\boldsymbol{x})<0,g(\boldsymbol{x})\geqslant\boldsymbol{0},\boldsymbol{h}(\boldsymbol{x})=\boldsymbol{0}$
2. 系统2：$\omega_0\phi(\boldsymbol{x})-\boldsymbol{\lambda}^T\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{\mu}^T\boldsymbol{h}(\boldsymbol{x})\geqslant 0,\forall\boldsymbol{x}\in\boldsymbol{D},(\omega_0,\boldsymbol{\lambda})\geqslant\boldsymbol{0},(\omega_0,\boldsymbol{\lambda},\boldsymbol{\mu})\neq\boldsymbol{0}$

强对偶定理的引理的证明1

设系统1无解，定义集合$C = \{(p, \boldsymbol{q}, \boldsymbol{r}) | \exists \boldsymbol{x} \in \boldsymbol{D}, p > \phi(\boldsymbol{x}), \boldsymbol{q} \leqslant \boldsymbol{g}(\boldsymbol{x}), \boldsymbol{r} = \boldsymbol{h}(\boldsymbol{x})\}$，先证明$C$是非空凸集。

以下证明C是非空凸集：由于D非空，所以C非空，任取$(p_1, \boldsymbol{q}_1, \boldsymbol{r}_1), (p_2, \boldsymbol{q}_2, \boldsymbol{r}_2) \in \boldsymbol{C}$，则存在$\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{D}$，使得$p_1 > \phi(\boldsymbol{x}_1), \boldsymbol{q}_1 \leqslant \boldsymbol{g}(\boldsymbol{x}_1), \boldsymbol{r}_1 = \boldsymbol{h}(\boldsymbol{x}_1), p_2 > \phi(\boldsymbol{x}_2), \boldsymbol{q}_2 \leqslant \boldsymbol{g}(\boldsymbol{x}_2), \boldsymbol{r}_2 = \boldsymbol{h}(\boldsymbol{x}_2)$。

对任意的$\lambda \in [0, 1]$，记$(\hat{p}, \hat{\boldsymbol{q}}, \hat{\boldsymbol{r}}) = \lambda(p_1, \boldsymbol{q}_1, \boldsymbol{r}_1) + (1 - \lambda)(p_2, \boldsymbol{q}_2, \boldsymbol{r}_2) = (\lambda p_1 + (1 - \lambda)p_2, \lambda \boldsymbol{q}_1 + (1 - \lambda)\boldsymbol{q}_2, \lambda \boldsymbol{r}_1 + (1 - \lambda)\boldsymbol{r}_2)$

由于$\phi(\boldsymbol{x})$是凸函数，$\boldsymbol{g}(\boldsymbol{x})$的每个分量是凹函数，$\boldsymbol{h}(\boldsymbol{x})$的每个分量是线性函数，所以存在$\hat{\boldsymbol{x}} = \lambda \boldsymbol{x}_1 + (1 - \lambda)\boldsymbol{x}_2 \in \boldsymbol{D}$：

$$\lambda p_1 + (1 - \lambda)p_2 > \lambda \phi(\boldsymbol{x}_1) + (1 - \lambda)\phi(\boldsymbol{x}_2) \geqslant \phi(\lambda \boldsymbol{x}_1 + (1 - \lambda)\boldsymbol{x}_2) = \phi(\hat{\boldsymbol{x}})$$
$$\lambda \boldsymbol{q}_1 + (1 - \lambda)\boldsymbol{q}_2 \leqslant \lambda \boldsymbol{g}(\boldsymbol{x}_1) + (1 - \lambda)\boldsymbol{g}(\boldsymbol{x}_2) \leqslant \boldsymbol{g}(\lambda \boldsymbol{x}_1 + (1 - \lambda)\boldsymbol{x}_2) = \boldsymbol{g}(\hat{\boldsymbol{x}})$$
$$\lambda \boldsymbol{r}_1 + (1 - \lambda)\boldsymbol{r}_2 = \lambda \boldsymbol{h}(\boldsymbol{x}_1) + (1 - \lambda)\boldsymbol{h}(\boldsymbol{x}_2) = \boldsymbol{h}(\lambda \boldsymbol{x}_1 + (1 - \lambda)\boldsymbol{x}_2) = \boldsymbol{h}(\hat{\boldsymbol{x}})$$

所以$(p, \boldsymbol{q}, \boldsymbol{r}) \in \boldsymbol{C}$，所以$C$是非空凸集。

强对偶定理的引理的证明1

如果系统1无解，则$(0,\mathbf{0},\mathbf{0}) \notin \boldsymbol{C}$，根据点与凸集的分离定理可知，存在$(\bar{\omega}_0,\bar{\boldsymbol{\lambda}},\bar{\boldsymbol{\mu}}) \neq \mathbf{0}$使得对于每一个$(p,\boldsymbol{q},\boldsymbol{r}) \in cl\boldsymbol{C}$都有$\bar{\omega}_0 p - \bar{\boldsymbol{\lambda}}^T\boldsymbol{q} - \bar{\boldsymbol{\mu}}\boldsymbol{r} \geqslant 0$，令$(\omega_0,-\boldsymbol{\lambda},-\boldsymbol{\mu}) = (\bar{\omega}_0,-\bar{\boldsymbol{\lambda}},-\bar{\boldsymbol{\mu}})$，可将上式写成$\omega_0 p - \boldsymbol{\lambda}^T\boldsymbol{q} - \boldsymbol{\mu}\boldsymbol{r} \geqslant 0, \forall (p,\boldsymbol{q},\boldsymbol{r}) \in cl\boldsymbol{C}$。固定$\boldsymbol{x} \in \boldsymbol{D}$，取$p,\boldsymbol{q},\boldsymbol{r}$，使得$p > \phi(\boldsymbol{x}), \boldsymbol{q} \leqslant \boldsymbol{g}(\boldsymbol{x}), \boldsymbol{r} = \boldsymbol{h}(\boldsymbol{x})$，在满足上述条件下，$p$可以取任意大的正数，$\boldsymbol{q}$的分量可以取任意小的负数，均有$(p,\boldsymbol{q},\boldsymbol{r}) \in \boldsymbol{C}$恒成立，所以必有$\omega_0 \geqslant 0, \boldsymbol{\lambda} \geqslant \mathbf{0}$。令$(p,\boldsymbol{q},\boldsymbol{r}) = (\phi(\boldsymbol{x}),\boldsymbol{g}(\boldsymbol{x}),\boldsymbol{h}(\boldsymbol{x})) \in cl\boldsymbol{C}$，从而可得$\omega_0\phi(\boldsymbol{x}) - \boldsymbol{\lambda}^T\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\mu}^T\boldsymbol{h}(\boldsymbol{x}) \geqslant 0, \forall \boldsymbol{x} \in \boldsymbol{D}, (\omega_0,\boldsymbol{\lambda}) \geqslant \mathbf{0}, (\omega_0,\boldsymbol{\lambda},\boldsymbol{\mu}) \neq \mathbf{0}$，即系统2存在解$(\omega_0,\boldsymbol{\lambda},\boldsymbol{\mu})$。

强对偶定理的引理的证明2

如果系统2存在解$(\omega_0, \boldsymbol{\lambda}, \boldsymbol{\mu})$，其中$\omega_0 > 0, \boldsymbol{\lambda} \geqslant \mathbf{0}$满足$\omega_0 \phi(\boldsymbol{x}) - \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\mu}^T \boldsymbol{h}(\boldsymbol{x}) \geqslant 0, \forall \boldsymbol{x} \in \boldsymbol{D}$。
假设存在$\boldsymbol{x} \in \boldsymbol{D}$，使得$\boldsymbol{g}(\boldsymbol{x}) \geqslant \mathbf{0}, \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}$，由于$\boldsymbol{\lambda} \geqslant \mathbf{0}$，则$\boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x}) \geqslant 0$，可得$\omega_0 \phi(\boldsymbol{x}) \geqslant 0$，由于$\omega_0 > 0$，所以可得$\phi(\boldsymbol{x}) \geqslant 0$，与系统1的要求有冲突，则系统1无解。

强对偶定理

假设

- 1）$\boldsymbol{D}$是非空凸集，$f(\boldsymbol{x})$是凸函数，$g_i(\boldsymbol{x}), i=1,2,\cdots,m$是凹函数，$h_j(\boldsymbol{x}), j=1,2,\cdots,l$为线性函数，即$\boldsymbol{h}(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}$。
- 2）假设存在$\hat{\boldsymbol{x}}\in\boldsymbol{D}$使得$g_i(\hat{\boldsymbol{x}})>0, i=1,2,\cdots,m; h_i(\hat{\boldsymbol{x}})=0, i=1,2,\cdots,l$且$\boldsymbol{0}\in int\ H(\boldsymbol{D})$，其中$H(\boldsymbol{D})=\{h_1(\boldsymbol{x}),h_2(\boldsymbol{x}),\cdots,h_l(\boldsymbol{x})|\boldsymbol{x}\in\boldsymbol{D}\}$
  （第二个假设条件也被称为Slater’s condition，int是指内点）

则强对偶成立，即
$inf\{f(\boldsymbol{x})|\boldsymbol{g}(\boldsymbol{x})\geqslant\boldsymbol{0},\boldsymbol{h}(\boldsymbol{x})=\boldsymbol{0},\boldsymbol{x}\in\boldsymbol{D}\}=sup\{\theta(\boldsymbol{\lambda},\boldsymbol{\mu})|\boldsymbol{\lambda}\geqslant\boldsymbol{0}\}$。
另外，如果inf为有限值，则$sup\{\theta(\boldsymbol{\lambda},\boldsymbol{\mu})|\boldsymbol{\lambda}\geqslant\boldsymbol{0}\}$在$(\bar{\omega},\bar{\mu})$达到，$\bar{\omega}\geqslant\boldsymbol{0}$。
如果 inf在点$\bar{\boldsymbol{x}}$达到，则$\bar{\boldsymbol{\lambda}}^T\boldsymbol{g}(\bar{\boldsymbol{x}})=0$。

强对偶定理的证明

证明

- 设$r = inf\{f(\boldsymbol{x})|g(\boldsymbol{x}) \geqslant \boldsymbol{0}, \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{x} \in \boldsymbol{D}\}.$，若$r = -\infty$，则由弱对偶定理的推论3可得$sup\{\theta(\boldsymbol{\lambda}, \boldsymbol{\mu})|\boldsymbol{\lambda} \geqslant \boldsymbol{0}\} = -\infty$，即强对偶成立
- 假设r是有限值，考虑系统：$f(\boldsymbol{x}) - r < 0, \boldsymbol{g}(\boldsymbol{x}) \geqslant \boldsymbol{0}, \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{x} \in \boldsymbol{D}$
- 由r的定义可知此系统无解，根据强对偶定理的引理可知存在$(\omega_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) \neq \boldsymbol{0}, (\omega_0, \boldsymbol{\lambda}) \geqslant \boldsymbol{0}$，使得对每个$\boldsymbol{x} \in \boldsymbol{D}$都有$\omega_0[f(\boldsymbol{x}) - r] - \boldsymbol{\lambda}^T\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\mu}\boldsymbol{h}(\boldsymbol{x}) \geqslant 0$
- 假设$\omega_0 = 0$，则$\boldsymbol{\lambda}^T\boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{\mu}^T\boldsymbol{h}(\boldsymbol{x}) \leqslant 0$，又因已知存在$\hat{\boldsymbol{x}} \in \boldsymbol{D}$使得$\boldsymbol{g}(\hat{\boldsymbol{x}}) > \boldsymbol{0}, \boldsymbol{h}(\hat{\boldsymbol{x}}) = \boldsymbol{0}$，代入可得$\boldsymbol{\lambda} = \boldsymbol{0}$，也即$\boldsymbol{\mu}^T\boldsymbol{h}(\hat{\boldsymbol{x}}) \leqslant \boldsymbol{0}, \forall \boldsymbol{x} \in \boldsymbol{D}$
- 由于$\boldsymbol{0} \in intH(\boldsymbol{D})$，因此可令$\boldsymbol{x} \in \boldsymbol{D}$，使得$\boldsymbol{h}(\boldsymbol{x}) = \lambda\boldsymbol{\mu}$其中$\lambda > 0$，于是有$0 \geqslant \boldsymbol{\mu}^T\boldsymbol{h}(\boldsymbol{x}) = \lambda||\boldsymbol{\mu}||^2$，于是$\boldsymbol{\mu} = \boldsymbol{0}$
- 因此$\omega_0 = 0 \rightarrow (\omega_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \boldsymbol{0}$，与假设不一致，所以$\omega_0 \neq 0$

强对偶定理的证明

- 由于$\omega_0 \neq 0$，可令$\bar{\boldsymbol{\lambda}} = \frac{\boldsymbol{\lambda}}{\omega_0}, \bar{\boldsymbol{\mu}} = \frac{\boldsymbol{\mu}}{\omega_0}$，所以$\omega_0[f(\boldsymbol{x}) - r] - \boldsymbol{\lambda}^T \boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{\mu}\boldsymbol{h}(\boldsymbol{x}) \geqslant 0$可转化为$f(\boldsymbol{x}) - \bar{\boldsymbol{\lambda}}^T \boldsymbol{g}(\boldsymbol{x}) - \bar{\boldsymbol{\mu}}\boldsymbol{h}(\boldsymbol{x}) \geqslant r$
- 可得$\theta(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) = inf\{f(\boldsymbol{x}) - \bar{\boldsymbol{\lambda}}^T \boldsymbol{g}(\boldsymbol{x}) - \bar{\boldsymbol{\mu}}^T \boldsymbol{h}(\boldsymbol{x}) | \boldsymbol{x} \in \boldsymbol{D}\} \geqslant r$，其中$\bar{\boldsymbol{\lambda}} \geqslant \boldsymbol{0}$
- 又由于弱对偶$f(\boldsymbol{x}) \geqslant \theta(\boldsymbol{\lambda}, \boldsymbol{\mu})$和$r$的定义$r = inf\{f(\boldsymbol{x}) | \boldsymbol{g}(\boldsymbol{x}) \geqslant \boldsymbol{0}, \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{x} \in \boldsymbol{D}\}$所以$\theta(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}}) = r$，$(\bar{\boldsymbol{\lambda}}, \bar{\boldsymbol{\mu}})$是对偶问题的最优解
- 此外，$\bar{\boldsymbol{x}}$是原问题的最优解，即满足$\bar{\boldsymbol{x}} \in \boldsymbol{D}, \boldsymbol{g}(\bar{\boldsymbol{x}}) \geqslant \boldsymbol{0}, \boldsymbol{h}(\bar{\boldsymbol{x}}) = \boldsymbol{0}, f(\bar{\boldsymbol{x}}) = r$，代入$f(\boldsymbol{x}) - \bar{\boldsymbol{\lambda}}^T \boldsymbol{g}(\boldsymbol{x}) - \bar{\boldsymbol{\mu}}\boldsymbol{h}(\boldsymbol{x}) \geqslant r$可得$\bar{\boldsymbol{\lambda}}^T \boldsymbol{g}(\bar{\boldsymbol{x}}) \leqslant 0$
- 由于$\bar{\boldsymbol{\lambda}} \geqslant \boldsymbol{0}, \boldsymbol{g}(\bar{\boldsymbol{x}}) \geqslant \bar{0}$可知$\bar{\boldsymbol{\lambda}}^T \boldsymbol{g}(\bar{\boldsymbol{x}}) = 0$，强对偶定理成立

对偶定理的应用-线性规划(1)

原问题：

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x}$$

$$s.t.\ \boldsymbol{A}\boldsymbol{x} \geqslant \boldsymbol{b}$$

其中$\boldsymbol{A}$是$m \times n$矩阵，$\boldsymbol{b}$是$n \times 1$向量

对偶问题：

$$max\ \theta(\boldsymbol{\lambda})$$

$$s.t. \boldsymbol{\lambda} \geqslant \boldsymbol{0}$$

其中

$$\begin{aligned}\theta(\boldsymbol{\lambda}) &= inf\{\boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{\lambda}^T(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})\} \\ &= inf\{(\boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{\lambda})^T\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{\lambda}\}\end{aligned}$$

对偶定理的应用-线性规划(1)

由于$\boldsymbol{x} \in \mathbb{R}^n$，所以

$$\theta(\boldsymbol{\lambda}) = \begin{cases} \boldsymbol{b}^T\boldsymbol{\lambda},\ if\ \boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{\lambda} = 0 \\ -\infty,\ otherwise \end{cases}$$

于是对偶问题可以重写为：

$$\max_{\boldsymbol{\lambda} \in \mathbb{R}^m}\ \boldsymbol{b}^T\boldsymbol{\lambda}$$

$$s.t.\ \boldsymbol{A}^T\boldsymbol{\lambda} = \boldsymbol{c}$$

$$\boldsymbol{\lambda} \geqslant \boldsymbol{0}$$

对偶定理的应用-线性规划(2)

原问题：

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x}$$

$$s.t.\ \boldsymbol{A}\boldsymbol{x} \geqslant \boldsymbol{b}, \boldsymbol{x} \geqslant \boldsymbol{0}$$

其中$\boldsymbol{A}$是$m\times n$矩阵，$\boldsymbol{b}$是$n\times 1$向量

对偶问题：

$$\max\ \theta(\boldsymbol{\lambda})$$

$$s.t. \boldsymbol{\lambda} \geqslant \boldsymbol{0}$$

其中

$$\begin{aligned}\theta(\boldsymbol{\lambda}) &= \inf\{\boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{\lambda}^T(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x})\}\\ &= \inf\{(\boldsymbol{c}-\boldsymbol{A}^T\boldsymbol{\lambda})^T\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{\lambda}\}\end{aligned}$$

对偶定理的应用-线性规划(2)

由于$\boldsymbol{x} \in \mathbb{R}^n$，所以

$$\theta(\boldsymbol{\lambda}) = \begin{cases} \boldsymbol{b}^T\boldsymbol{\lambda}, \textit{if } \boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{\lambda} \geqslant 0 \\ -\infty, \textit{otherwise} \end{cases}$$

于是对偶问题可以重写为：

$$\max_{\boldsymbol{\lambda}\in\mathbb{R}^m} \boldsymbol{b}^T\boldsymbol{\lambda}$$

$$s.t. \ \boldsymbol{A}^T\boldsymbol{\lambda} \leqslant \boldsymbol{c}$$

$$\boldsymbol{\lambda} \geqslant \boldsymbol{0}$$

对偶定理的应用-二阶问题最小化

$$min\ \boldsymbol{x}^T\boldsymbol{x}$$

$$s.t.\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

对偶问题：

$$min\ \boldsymbol{x}^T\boldsymbol{x}$$

$$s.t.\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\theta(\boldsymbol{\mu}) = inf\{\boldsymbol{x}^T\boldsymbol{x} - \boldsymbol{\mu}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})\}$$

可以利用梯度算法令inf中的数值为0来求解最小值，可得$2\boldsymbol{x} - \boldsymbol{A}^T\boldsymbol{\mu} = \boldsymbol{0}$即$\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{\mu}/2$ 所以$\theta(\boldsymbol{\mu}) = -\frac{1}{4}\boldsymbol{\mu}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{b}^T\boldsymbol{\mu}$
对偶问题为$\max\limits_{\boldsymbol{\mu}\in\mathbb{R}^m} \ -\frac{1}{4}\boldsymbol{\mu}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{b}^T\boldsymbol{\mu}$

Thank you for your attention!

翟卫欣 副教授

zhaiweixin@cau.edu.cn

中国农业大学信息与电气工程学院