应用运筹学基础：线性规划 (4) - 对偶与对偶单纯形法

这一节课讲解了线性规划的对偶问题及其性质。

引入对偶问题

考虑一个线性规划问题：
$$\begin{aligned} \max_x \quad & 4x_1 + 3x_2 \\ \text{s.t.} \quad & 2x_1 + 3x_2 \le 24 \\ & 5x_1 + 2x_2 \le 26 \\ & x \ge 0 \end{aligned}$$
我们可以把这个问题看作一个生产模型：一份产品 A 可以获利 4 单位价格，生产一份需要 2 单位原料 C 和 5 单位原料 D；一份产品 B 可以获利 3 单位价格，生产一份需要 3 单位原料 C 和 2 单位原料 D。现有 24 单位原料 C，26 单位原料 D，问如何分配生产方式才能让获利最大。

但假如现在我们不生产产品，而是要把原料都卖掉。设 1 单位原料 C 的价格为 y_1，1 单位原料 D 的价格为 y_2，每种原料制定怎样的价格才合理呢？

首先，原料的价格应该不低于产出的产品价格（不然还不如自己生产...），所以我们有如下限制：
$$\begin{aligned} 2y_1 + 5y_2 \ge 4 \\ 3y_1 + 2y_2 \ge 3 \end{aligned}$$
当然也不能漫天要价（也要保护消费者利益嘛- -），所以我们制定如下目标函数：$\min_y \quad 24y_1 + 26y_2$ 合起来就是下面这个线性规划问题：
$$\begin{aligned} \min_y \quad & 24y_1 + 26y_2 \\ \text{s.t.} \quad & 2y_1 + 5y_2 \ge 4 \\ & 3y_1 + 2y_2 \ge 3 \\ & y \ge 0 \end{aligned}$$
这个问题就是原问题的对偶问题。

对偶问题

对于一个线性规划问题（称为原问题，primal，记为 P）$\begin{aligned}\max_x \quad & c^T x \\ \text{s.t.} \quad & Ax \le b \\ & x \ge 0\end{aligned}$ 我们定义它的对偶问题（dual，记为 D）为 $\begin{aligned}\min_x \quad & b^T y \\ \text{s.t.} \quad & A^T y \ge c \\ & y \ge 0\end{aligned}$ 这里的对偶变量 y，可以看作是对原问题的每个限制，都用一个变量来表示。

原问题限制条件的不等号，和对偶问题限制条件的不等号，是相互关联的。假设原问题是一个最大化问题，设 a_i^T 表示 A 中的第 i 行，我们有以下结论：

1. 若限制条件为 $a_i^T x \le b_i$，那么对偶问题中有 $y_i \ge 0$

证明略，根据对偶问题的定义即可获得。

2. 若限制条件为 $a_i^T x \ge b_i$，那么对偶问题中有 $y_i \le 0$

把不等式转换为标准形式显然有 $-a_i^T x \le -b_i$。令 $\bar{y}_i = -y_i$，用 $\bar{y}_i$ 表示原问题的第 i 个限制，令 $y' = [y_1 \quad \dots \quad y_{i-1} \quad \bar{y}_i \quad y_{i+1} \quad \dots \quad y_m]^T$，那么对偶问题可以写为

$$\begin{aligned}\min_{y'} \quad & [b_1 \quad \dots \quad b_{i-1} \quad -b_i \quad b_{i+1} \quad \dots \quad b_m] y' \\ \text{s.t.} \quad & [a_1 \quad \dots \quad a_{i-1} \quad -a_i \quad a_{i+1} \quad \dots \quad a_m] y' \ge c \\ & y' \ge 0\end{aligned}$$

将 $y_i = -\bar{y}_i$ 代回式中即可获得。

3. 若限制条件为 $a_i^T x = b_i$，那么对偶问题中对 y_i 无限制

$a_i^T x = b_i$ 可以看作 $a_i^T x \ge b_i$ 与 $a_i^T x \le b_i$，用 $\bar{y}_i$ 和 $\tilde{y}_i$ 表示这两个限制，并令 $y' = [y_1 \quad \dots \quad \bar{y}_i \quad \tilde{y}_i \quad \dots \quad y_m]^T$，那么对偶问题可以写为

$$\begin{aligned}\min_{y'} \quad & [b_1 \quad \dots \quad b_i \quad -b_i \quad \dots \quad b_m] y' \\ \text{s.t.} \quad & [a_1 \quad \dots \quad a_i \quad -a_i \quad \dots \quad a_m] y' \ge c \\ & y' \ge 0\end{aligned}$$

令 $y_i = \bar{y}_i - \tilde{y}_i$，代回上面的式子中即可获得原来的对偶问题的形式。容易看出 y_i 是可正可负的，没有限制。

4. 若 $x_i \geq 0$，那么对偶问题中有 $A_i^T y \geq c_i$ 5. 若 $x_i \leq 0$，那么对偶问题中有 $A_i^T y \leq c_i$ 6. 若 x_i 无限制，那么对偶问题中有 $A_i^T y = c_i$

这三条的推导和前三条类似，这里不再赘述。

对偶问题的性质

这一部分讲解线性规划中对偶问题的若干性质。

对称性

P 的对偶是 D，那么 D 的对偶也是 P。如果我们把对偶问题变成标准形式，有

$$\begin{aligned} \max_y \quad & y^T(-b) \\ \text{s.t.} \quad & (-A^T)y \leq -c \\ & y \geq 0 \end{aligned}$$

它的对偶问题是

$$\begin{aligned} \min_x \quad & -c^T x \\ \text{s.t.} \quad & -Ax \geq -b \\ & x \geq 0 \end{aligned}$$

把目标函数和限制都乘以 -1 之后就是原问题。

弱对偶定理 (weak duality)

设 x 和 y 分别是原问题和对偶问题的可行解，我们有 $c^T x \leq b^T y$。这是因为，由 y 的可行性我们有 $A^T y \geq c$，即 $y^T A \geq c^T$，两边同乘以 x 有 $y^T A x \geq c^T x$；由 x 的可行性我们还有 $Ax \leq b$，那么 $y^T A x \leq y^T b$，合起来就是 $c^T x \leq b^T y$。

由弱对偶定理我们马上获得以下两条性质。

最优性

若 x 和 y 分别是原问题和对偶问题的可行解，而且 $c^T x = b^T y$，那么 x 和 y 分别是原问题和对偶问题的最优解。

无界性

若原问题有可行解无最优解（就是目标函数值可以取无穷大），那么对偶问题无可行解；若对偶问题有可行解无最优解，那么原问题无可行解。

当然啦，也有两个问题都无解的情况发生，比如下面这个线性规划 $\begin{aligned}\max_x \quad & x_1 + x_2 \\ \text{s.t.} \quad & x_1 - x_2 \leq 1 \\ & -x_1 + x_2 \leq -2 \\ & x \geq 0\end{aligned}$ 它的对偶问题是 $\begin{aligned}\min_y \quad & y_1 - 2y_2 \\ \text{s.t.} \quad & y_1 - y_2 \geq 1 \\ & -y_1 + y_2 \geq 1 \\ & y \geq 0\end{aligned}$

强对偶定理 (strong duality)

若原问题（或对偶问题）有有限最优解，那么对偶问题（或原问题）也有有限最优解，且二者最优解相等。

可以通过单纯形法的计算过程来辅助证明。

假设引入松弛变量之后，原问题变为 $\begin{aligned}\max_x \quad & c^T x \\ \text{s.t.} \quad & \bar{A}x = b \\ & x \geq 0\end{aligned}$ 画出初始的单纯形表 $\begin{array}{c|cc|c} & c^T & 0 & 0 \\ \hline x_B^* & A & I & b\end{array}$ 最终的单纯形表为 $\begin{array}{c|cc|c} & c^T - c_B^T \bar{A}_B^{-1} A & -c_B^T \bar{A}_B^{-1} & -c_B^T \bar{A}_B^{-1} b \\ \hline x_B^* & \bar{A}_B^{-1} A & \bar{A}_B^{-1} & \bar{A}_B^{-1} b\end{array}$ 由于 x_B^* 是原问题最优的基变量组合，那么检验数均非正，即 $\begin{aligned}c^T & \leq c_B^T \bar{A}_B^{-1} A \\ -c_B^T \bar{A}_B^{-1} & \leq 0\end{aligned}$ 不妨取 $y^{*T} = c_B^T \bar{A}_B^{-1}$，根据上面两个式子，我们有 $A^T y^* \geq c$ 以及 $y^* \geq 0$，即 y^* 是一个可行解。根据单纯形法，我们知道 $x_B^* = \bar{A}_B^{-1} b$，$x_N^* = 0$，那么对偶问题的目标函数值为 $b^T y^* = y^{*T} b = c_B^T \bar{A}_B^{-1} b = c_B^T x_B^* = c^T x^*$，即我们找到了一对可行的 x 和 y，使得原问题和对偶问题的目标函数值相等，那么根据弱对偶定理，这两个可行解分别是原问题和对偶问题的最优解。

互补松弛定理 (complementary slackness)

若 x^* 与 y^* 分别是原问题和对偶问题的可行解，那么以下两点等价：

\1. x^* 和 y^* 分别是原问题和对偶问题的最优解；

\2. $(y^{*T}A - c^T)x^* = 0$ 且 $y^{*T}(Ax^* - b) = 0$。

由 2 推出 1 很简单，把括号都打开后有 $y^{*T}b = y^{*T}Ax^* = c^Tx^*$，根据弱对偶定理得 x^* 和 y^* 分别是原问题和对偶问题的最优解。

由 1 推出 2 也不难，根据弱对偶定理中的推导，我们有 $y^{*T}b \geq y^{*T}Ax^* \geq c^Tx^*$。而 x^* 和 y^* 分别是原问题和对偶问题的最优解，那么 $y^{*T}b = c^Tx^*$，不等式就会全部取等，即 $y^{*T}b = y^{*T}Ax^* = c^Tx^*$，加上括号就行了。

这个定理揭示了原始问题的最优解和对偶问题的最优解之间的关系，它们对限制条件的满足是“一紧一松”的。

对偶单纯形法

利用强对偶定理，我们可以为单纯形法作出另一种解释。

我们知道，单纯形法的停止条件是所有检验数非正（如果是 min 问题就是所有检验数非负）。而从强对偶定理的推导中我们可以看到，所有检验数非正时，我们就能构造一个对偶问题的可行解，使得原问题和对偶问题的目标函数值相等，那么它们分别是原问题和对偶问题的最优解。也就是说，单纯形法是在保证原问题可行解的情况下，尝试构造对偶问题的可行解（这个可行解让目标函数值与原问题相同），如果构造成功，那么两个都是最优解。这种单纯形法又称为原始单纯形法。

相应地，我们可以设计对偶单纯形法：在保证对偶问题可行解（所有检验数非正，如果是 min 问题就是所有检验数非负）的情况下，尝试构造原始问题的可行解（这个可行解让目标函数值与对偶问题相同），如果构造成功，那么两个都是最优解。

下面以 min 问题为例，简要说明对偶单纯形法的计算步骤：

\1. 找到一组基，使得所有检验数非负；

\2. 如果单纯形表中 b 的那一列出现负数，说明当前基不可行（因为有 $x \ge 0$ 的限制），选择负数中 b 的绝对值最大的那一行（设为第 i 行），对应的变量 x_i 作为出基变量（要把该变量从负数调到 0）；

\3. 假设第 i 行中，x_j 的系数为 a_j，检验数为 d_j，那么在所有 $a_j < 0$ 的变量中，选择 d_j/a_j 绝对值最小的那一列，对应的变量 x_k 作为入基变量，回到 2。如果所有 $a_j \ge 0$，那么原问题无可行解；

\4. 如果单纯形表中 b 的那一列均非负，说明构造出了一个原问题的可行解，算法结束。

为什么要选择 $a_k < 0$ 的变量呢？我们写出出基变量和非基变量之间的关系式 $\sum_{j\in N} a_j x_j + x_i = b_i$ 如果 $a_k > 0$，那么为了把 x_i 从负数调到 0，又要保证等式成立，x_k 只能从 0 变成负数，就不能入基了；相反，如果 $a_k < 0$，那么为了让等式成立，x_k 会从 0 变成正数，就可以入基，向原始问题的可行解靠近一步。

为什么要选择 d_j/a_j 绝对值最小的变量呢？我们写出 x_k 和其它变量的关系式，以及目标函数和检验数的关系式：

$$x_k = \frac{b_i - x_i}{a_k} - \sum_{j\in N, j\neq k} \frac{a_j}{a_k} x_j$$

$$z = v + \sum_{j\in N} d_j x_j = \sum_{j\in N, j\neq k} (d_j - \frac{d_k}{a_k} a_j) x_j - \frac{d_k}{a_k} x_i + (v + \frac{d_k}{a_k} b_i)$$

为了保持对偶问题的可行解，我们需要保证变量替换之后，检验数仍然非负，即

$$-\frac{d_k}{a_k} \ge 0$$

$$d_j - \frac{d_k}{a_k} a_j \ge 0$$

第一个式子显然满足，因为原检验数 $d_k \ge 0$，且 $a_k < 0$。第二个式子在 a_j 为正数时显然满足，a_j 为负数时，需要 $\frac{d_j}{a_j} \le \frac{d_k}{a_k}$ 才能满足，这就是选择 d_j/a_j 绝对值最小的变量的原因。

举一个例子 $\begin{aligned} \min_x \quad & 9x_1+5x_2+3x_3 \\ \text{s.t.} \quad & 3x_1+2x_2-3x_3 \geq 3 \\ & 2x_1+x_3 \geq 5 \\ & x \geq 0 \end{aligned}$ 加入松弛变量后，问题转化为

$$\begin{aligned} \min_x \quad & 9x_1+5x_2+3x_3 \\ \text{s.t.} \quad & 3x_1+2x_2-3x_3-x_4 = 3 \\ & 2x_1+x_3-x_5 = 5 \\ & x \geq 0 \end{aligned}$$

绘制单纯形表，第一次迭代：

	9	5	3	0	0	0
x_4	-3	-2	3	1	0	-3
x_5	-2	0	-1	0	1	-5

选择 x_5 出基. 由于 $3/1 < 9/2$，选择 x_3 入基，第二次迭代：

	3	5	0	0	3	-15
x_4	-9	-2	0	1	3	-18
x_3	2	0	1	0	-1	5

选择 x_4 出基. 由于 $3/9 < 5/2$，选择 x_1 入基，第三次迭代：

	0	$13/3$	0	$1/3$	4	-21
x_1	1	$2/9$	0	$-1/9$	$-1/3$	2
x_3	0	$-4/9$	1	$2/9$	$-1/3$	1

此时最后一列第二、三行均非负，迭代结束。原问题的最优解为 $x_1=2, x_2=0, x_3=1$，目标函数值为 21。