对偶理论

针对任意一个最优化问题，可以定义它的一个对偶问题 (Dual Problem) 。

对偶理论将揭示原问题与对偶问题之间的内在联系，为进一步深入研究线性规划的求解算法提供理论依据。

对偶理论

原问题：

$$\begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}. \end{array} \tag{LP}$$

定义Lagrange函数 $L(x, \lambda, \mu) = c^\top x + \lambda^\top (b - Ax) - \mu^\top x, \mu \geq 0, \lambda \in \mathbb{R}^m$. 对于无约束问题

$$\min_{x \in \mathbb{R}^n} L(x, \lambda, \mu),$$

我们记 $g(\lambda, \mu) = \min_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$.

对于任意原问题的可行解$\bar{x}$，我们有

$$g(\lambda, \mu) \leq c^\top \bar{x} + \lambda^\top (b - A\bar{x}) - \mu^\top x \leq c^\top \bar{x}.$$

故，问题

$$\max_{\lambda, \mu} g(\lambda, \mu)$$

可以看作寻找原问题的最紧下界。

$$\begin{aligned} \min_{x \in \mathbb{R}^n} L(x, \lambda, \mu) &= \min_x (c - A^\top \lambda - \mu)^\top x + \lambda^\top b \\ &= \begin{cases} b^\top \lambda & \text{if } A^\top \lambda + \mu = c, \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

对偶问题：

$$\begin{array}{ll} \max & \mathbf{b}^T\lambda \\ \text{s.t.} & A^T\lambda \leq \mathbf{c}. \end{array} \tag{DP}$$

对偶理论

$$\begin{aligned} &\min f(x) \\ \text{s.t.} \quad & c_i(x) = 0, i \in \mathcal{E}, \\ & c_i(x) \geq 0, i \in \mathcal{I}. \end{aligned}$$

一般优化问题的Langrange 对偶：

$$\max_{\lambda,\mu \geq 0} \min_x L(x, \lambda, \mu) = f(x) + \sum_{i \in \mathcal{E}} \lambda_i c_i(x) - \sum_{i \in \mathcal{I}} \mu_i c_i(x)$$

习题：2.5 推导出下面问题的对偶问题：

1. $$\min c^\top x, \quad \text{s.t. } Ax \geq b.$$
2. $$\min c^\top x, \quad \text{s.t. } Ax \leq b.$$
3. 并说明：(DP)的对偶等价于(LP). 即，对偶的对偶是原问题。

对偶理论

定理

弱对偶定理： 任意线性规划问题(LP)及其对偶问题(DP)之间成立以下关系

$\mathbf{x}$是(LP)问题的可行解，λ是(DP)问题的可行解， $\Rightarrow \mathbf{c}^T\mathbf{x} \geq \mathbf{b}^T\lambda.$

推论：

(a) $\mathbf{x}$是(LP)问题的可行解 $\Rightarrow \mathbf{c}^T\mathbf{x} \geq$ 问题(DP)的最大值。
λ是(DP)问题的可行解 $\Rightarrow \mathbf{b}^T\lambda \leq$ 问题(LP)的最小值。

(b) $\mathbf{x}$是(LP)问题的可行解，λ是(DP)问题的可行解，且$\mathbf{c}^T\mathbf{x} = \mathbf{b}^T\lambda$ $\Rightarrow$ $\mathbf{x}$是(LP)问题的最优解，λ是(DP)问题的最优解。

(c) 问题(LP)无界 $\Rightarrow$ 问题(DP)无可行解。
问题(DP)无界 $\Rightarrow$ 问题(LP)无可行解。

(d) **最优解存在性定理：** 若(LP)与(DP)都有可行解，则它们均存在最优解。

定理

强对偶定理： 原问题(LP) 有最优解，则对偶问题(DP)也有最优解，且此时两方的最优值一致。

Proof.

设 (LP)的最优解为某个可行基解： $\bar{x} = (x_B, 0)^\top$, 那么最优值为 $c_B^\top x_B$. 由最优判定定理得：

$$c^\top - c_B^\top B^{-1} A \geq 0^\top.$$

令$\bar{\lambda} = B^{-T} c_B \in \mathbb{R}^m$, 那么，

$$A^\top \bar{\lambda} \leq c.$$

并且

$$\bar{\lambda}^\top b = c_B^\top B^{-1} b = c_B^\top x_B.$$

即, $\bar{\lambda}$是(DP)的最优解，且两者最优值一致。 □

对偶理论

互补松弛定理： 设$\bar{\mathrm{x}}$和$\bar{\lambda}$分别是(LP)和(DP)的可行解，那么$\bar{\mathrm{x}}$和$\bar{\lambda}$是对应问题最优解的充要条件是：

$$\begin{cases} \bar{\mathbf{x}}^T(A^T\bar{\lambda} - \mathbf{c}) = 0 \\ \bar{\lambda}^T(A\bar{\mathbf{x}} - \mathbf{b}) = 0 \end{cases} \tag{27}$$

回顾：线性规划的KKT条件为：

$$\text{(KKT)} \qquad \begin{cases} A\mathbf{x} - \mathbf{b} = 0 \\ \mathbf{x} \geq \mathbf{0} \\ A^T\lambda - \mathbf{c} \leq 0 \\ \mathbf{x}^T(A^T\lambda - \mathbf{c}) = 0 \\ \lambda^T(A\mathbf{x} - \mathbf{b}) = 0 \end{cases} \tag{28}$$

KKT条件为充分条件： Slater's condition。与此处的结论一致。

利用互补松弛定理，当知道一个问题的最优解时，可求出其对偶问题的最优解。

$$\begin{array}{ll} \min & 13x_1+10x_2+6x_3 \\ \text{s.t.} & 5x_1+x_2+3x_3=8 \\ & 3x_1+x_2=3 \\ & x_1\geq 0, x_2\geq 0, x_3\geq 0. \end{array}$$

的最优解$\bar{\mathbf{x}}=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

$$\begin{array}{ll} \max & 8\lambda_1+3\lambda_2 \\ \text{s.t.} & 5\lambda_1+3\lambda_2\leq 13 \\ & \lambda_1+\lambda_2\leq 10 \\ & 3\lambda_1\leq 6. \end{array}$$

的最优解$\bar{\lambda}=\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}=\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

灵敏度分析

实际问题中，很多时候是基于某些采集数据来决定模型的系数。在这种情况下，势必会出现系数的扰动及引起的变化。

所谓灵敏度分析就是利用解一个问题得到的结果，研究当系数有微小变化时最优解的反应。

灵敏度分析

为简单起见，考虑问题

$$\begin{aligned} \min \quad & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{29}$$

约束条件右边常数向量作微小变化

$$\mathbf{b} + \Delta\mathbf{b} = (b_1 + \Delta b_1, \cdots, b_m + \Delta b_m)^T,$$

得到新的线性规划问题

$$\begin{aligned} \min \quad & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} + \Delta\mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned} \tag{30}$$

灵敏度分析

现假设对问题(29)利用单纯形法已得到其最优解 $\mathbf{x}^* = \begin{pmatrix} \mathbf{x}_B^* \\ \mathbf{x}_N^* \end{pmatrix} = \begin{pmatrix} B^{-1}\mathbf{b} \\ \mathbf{0} \end{pmatrix}$.

这时对于最优基B, 根据可行性条件有

$$B^{-1}\mathbf{b} \geq \mathbf{0}, \tag{31}$$

根据最优性条件有

$$\mathbf{c}_N^T - \mathbf{c}_B^T B^{-1} N \geq \mathbf{0}. \tag{32}$$

当$\mathbf{b}$变为$\mathbf{b} + \Delta\mathbf{b}$后，如果还成立

$$B^{-1}(\mathbf{b} + \Delta\mathbf{b}) \geq \mathbf{0}, \tag{33}$$

则问题(29)的最优基B也是问题(30)的最优基。

此时问题(30)的最优解成为 $\mathbf{x} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} B^{-1}(\mathbf{b} + \Delta\mathbf{b}) \\ \mathbf{0} \end{pmatrix}$.

灵敏度分析

进一步，记问题(29)的目标函数最优值为$z(\mathbf{b})$, 问题(30)的目标函数最优值为$z(\mathbf{b}+\Delta\mathbf{b})$, 则有

$$\begin{aligned} z(\mathbf{b}) &= \mathbf{c}_B^T B^{-1}\mathbf{b} \\ z(\mathbf{b}+\Delta\mathbf{b}) &= \mathbf{c}_B^T B^{-1}(\mathbf{b}+\Delta\mathbf{b}) \end{aligned} \tag{34}$$

故成立

$$z(\mathbf{b}+\Delta\mathbf{b}) - z(\mathbf{b}) = \mathbf{c}_B^T B^{-1}\Delta\mathbf{b}. \tag{35}$$

令$(B^T)^{-1}\mathbf{c}_B = \lambda$. 根据对偶理论，它即为问题(29)的对偶问题的最优解。

因此，记$\lambda^* = (B^T)^{-1}\mathbf{c}_B$, 则上式可写成

$$z(\mathbf{b}+\Delta\mathbf{b}) - z(\mathbf{b}) = (\lambda^*)^T\Delta\mathbf{b} = \sum_{i=1}^m \lambda_i^*\Delta b_i. \tag{36}$$

考虑$\Delta b_i \to 0$时的极限，就可以得到

$$\frac{\partial z(\mathbf{b})}{\partial b_i} = \lambda_i^*, (i = 1, \cdots, m). \tag{37}$$

这里，实际上给出了对偶问题最优解的经济意义（影子价格）!

影子价格

例子来源 盒马鲜生的购物车界面处，可以推荐一些易逝品，这些商品当天不卖完，第二天就要报废或做折价处理。每个商品的库存数量是有限的。最理想的状态，应该是当天所有商品都能卖完。那每个用户进入购物车页面时，系统应该推荐哪些商品呢？

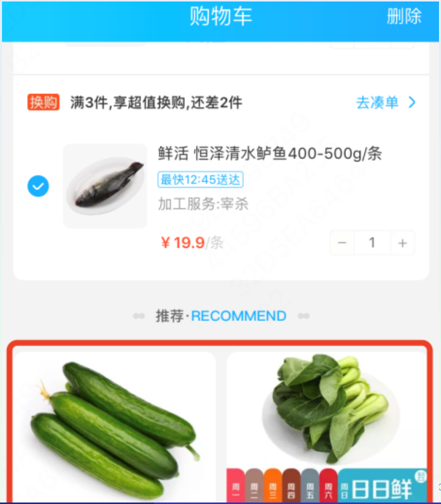
- 计算每件商品被推荐时能带来的价值

$$\mathrm{ctr}_i \times \mathrm{cvr}_i \times r_i,$$

  $\mathrm{ctr}_i, \mathrm{cvr}_i, r_i$
  分别为第件商品的点击率、转化率和价格。
- 有了该值后，
  将商品按照此值从大到小排列并推荐top K即可。
- 这个策略的一个潜在问题是未考虑商品的
  库存，可能会导致一些高价值的商品早早售罄，而中等价值的商品却因为推荐次数过少等原因没有卖完。
- 引入影子价格对上述策略进行微调：

$$\mathrm{ctr}_i \times \mathrm{cvr}_i \times (r_i - \lambda_i),$$

  λ_i为商品影子价格。
- 加了影子价格参数后，当高价
  值商品的库存充足时，商品的影子价格为0，不影响原推荐策略；当高价值商品的库存变少后，商品变得稀缺，影子价格便会大于0，从而降低高价值商品的排序，从而让位于其他价值稍低但库存较多的商品。



线性规划的多项式时间算法

从最坏计算量角度来看，单纯形方法不是多项式时间算法。关于单纯形方法的理论及实际上的计算效率可见综述文献：

R. Shamir. The efficiency of the simplex method: A survey. Management Science, Vol. 33, 1987: 301-334.

线性规划的多项式时间算法

针对线性规划问题的多项式时间算法：

- 椭球算法，最先在理论上明确线性规划问题恒在多项式时间内可解，因而它是意义深远的。
  L.G. Khachiyan. A polynomial algorithm in linear programming. Soviet Mathematics Doklady, Vol. 20, 1979: 191-194.
- 内点算法，作为可替代单纯形法的有效方法而引入注目，同时它可被推广并应用到不限于线性规划的更一般问题。
  N. Karmarkar. A new polynomial time algorithm for linear programming. Combinatorica, Vol. 4, 1984: 373-395.

线性规划求解器

- Gurobi: (商业求解器) https://www.gurobi.com/
- Google: ORtools: https://developers.google.com/optimization?hl=zh-cn
- CVX:(学术开源优化求解器) http://cvxr.com/cvx/doc/quickstart.html
- 国内公司：阿里巴巴(MindOpt:https://tianchi.aliyun.com/mindopt)、华为（天筹https://www.huaweicloud.com/product/modelarts/optverse.html)、杉数（https://www.shanshu.ai/)

本章作业（线性规划）提交截止时间：2023年9月29日0：00

- 习题2.1 证明：对于标准形式，如果矩阵$A \in \mathrm{R}^{m\times n}$的秩为$k$, $k < m$, 其行向量为 $a_1^\top, a_2^\top, \dots, a_m^\top$. 那么$A$的$k$个线性无关的行向量$a_{i_1}^\top, a_{i_2}^\top, \dots, a_{i_k}^\top$ 组成的子矩阵 $\tilde{A}$和对应的$\tilde{b}$, 有$P = Q$, 这里 $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$, $Q = \{x \mid \tilde{A}x = \tilde{b}, x \geq 0\}$.
- 习题2.2，证明 x是可行基解，则x是顶点。提示：构造$c = \sum_{i\in I} a_i$. I是主动集。
- 习题2.3：试给出下列问题的所有极点和基解

$$\begin{array}{ll} \min & -x_1 + 3x_2 \\ \text{s.t.} & x_1 + 2x_2 \leq 8 \\ & x_2 \leq 2 \\ & x_1, x_2 \geq 0. \end{array} \tag{38}$$

- 习题2.4：用单纯形法求解如下问题：

$$\begin{array}{ll} \min & -4x_1 - x_2 \\ \text{s.t.} & -x_1 + 2x_2 \leq 4 \\ & 2x_1 + 3x_2 \leq 12 \\ & x_1 - x_2 \leq 3 \\ & x_1, x_2 \geq 0. \end{array} \tag{39}$$

- 习题：2.5 推导出下面问题的对偶问题：
  1. $$\min c^\top x, \quad \text{s.t. } Ax \geq b.$$
  2. $$\min c^\top x, \quad \text{s.t. } Ax \leq b.$$
  3. 并说明：(DP)的对偶等价于(LP). 即，对偶的对偶是原问题。