Assignment

December 23, 2015

请将作业发送至：ggu@nju.edu.cn

- 邮件标题：or_姓名_学号
- 邮件内容：参照期末答案（期末考试后公布），自己预测的总评成绩
  - 总评=平时×0.2+期中×0.3+期末×0.5
  - 平时成绩按满分算
  - 邮件内容只允许出现一个 2 位数
- 邮件附件
  - 作业：姓名_学号.pdf
  - 程序：ipm_学号.m
  - 照片：姓名_学号.jpg

截止时间：2016 年 1 月 6 日晚 24 点之前

Que 1. Birkhoff's theorem (p3-30)
An $n \times n$ matrix X is called doubly stochastic if

$$X_{ij} \geq 0, \quad i,j = 1,\dots,n, \quad \sum_{i=1}^{n} X_{ij} = 1, \quad j = 1,\dots,n \quad \sum_{j=1}^{n} X_{ij} = 1, \quad i = 1,\dots,n.$$

In words, the entries of X are nonnegative, and its row and column sums are equal to one. The set of doubly stochastic matrices can be described as a polyhedron in $\mathbb{R}^{n^2}$, defined as

$$P = \{x \in \mathbb{R}^{n^2} \mid Cx = d, x \geq 0\}$$

with x the matrix X stored as a vector in column-major order,

$$\begin{aligned} x &= \mathrm{vec}(X) \\ &= (X_{11}, X_{21}, \dots, X_{n1}, X_{12}, X_{22}, \dots, X_{n2}, \dots, X_{1n}, X_{2n}, \dots, Xnn), \end{aligned}$$

and C, d defined as

$$C = \begin{bmatrix} I & I & \cdots & I \\ 1^T & 0 & \cdots & 0 \\ 0 & 1^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1^T \end{bmatrix}, \qquad d = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

(The identity matrices I have order n and the vectors 1 have length n.) The matrix C has size $2n \times n^2$ and the vector d has length $2n$. In this exercise we show that the extreme points of the set of doubly stochastic matrices are the permutation matrices.

(a) A permutation matrix is a $0-1$ matrix with exactly one element equal to one in each column and exactly one element equal to one in each row. Use the rank criterion for extreme points to show that all permutation matrices are extreme points of the polyhedron of doubly stochastic matrices. (More precisely, if X is a permutation matrix, then $\mathrm{vec}(X)$ is an extreme point of the polyhedron P of vectorized doubly stochastic matrices defined above.)

(b) Show that an extreme point X of the polyhedron of $n \times n$ doubly stochastic matrices has at most $2n-1$ nonzero entries. Therefore, if X is an extreme point, it must has a row with exactly one nonzero element (with value 1) and a column with exactly one nonzero element (equal to 1). Use this observation to show that all extreme points are permutation matrices.

Que 2. Goldman-Tucker Theorem (p6-25)
We consider an LP

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \leq b, \end{array}$$

with $A \in \mathbb{R}^{m \times n}$, and its dual

$$\begin{array}{ll} \max & -b^T z \\ \text{s.t.} & A^T z + c = 0, \quad z \geq 0. \end{array}$$

We assume the optimal value is finite. From duality theory we know that any primal optimal x^* and any dual optimal z^* satisfy the complementary slackness conditions

$$z_i^*(b_i - a_i^T x^*) = 0, \qquad i = 1, \ldots, m.$$

In other words, for each i, we have $z_i^* = 0$, or $a_i^T x^* = b_i$, or both.
In this problem you are asked to show that there exists at least one primal-dual optimal pair x^*, z^* that satisfies

$$z_i^*(b_i - a_i^T x^*) = 0, \qquad z_i^* + (b_i - a_i^T x^*) > 0, \qquad \text{for all } i.$$

This is called a strictly complementary pair. In a strictly complementary pair, we have for each i, $z_i^* = 0$, or $a_i^T x^* = b_i$, but not both.
To prove the result, suppose x^*, z^* are optimal but not strictly complementary, and

$$\begin{array}{lll} a_i^T x^* = b_i, & z_i^* = 0, & i = 1, \ldots, M \\ a_i^T x^* = b_i, & z_i^* > 0, & i = M+1, \ldots, N \\ a_i^T x^* < b_i, & z_i^* = 0, & i = N+1, \ldots, m \end{array}$$

with $M \geq 1$. In other words, $m - M$ entries of $b - Ax^*$ and z^* are strictly complementary; for the other entries we have zero in both vectors.

(a) Use Farkas' lemma to show that the following two sets of inequalities/equalities are strong alternatives:

- There exists a $v \in \mathbb{R}^n$ such that

$$\begin{aligned} a_1^T v &< 0 \\ a_i^T v &\leq 0, \quad i = 2, \ldots, M \\ a_i^T v &= 0, \quad i = M+1, \ldots N. \end{aligned} \tag{A1}$$

- There exists a $w \in \mathbb{R}^{N-1}$ such that

$$a_1 + \sum_{i=1}^{N-1} w_i a_{i+1} = 0, \qquad , w_i \geq 0, \quad i = 1, \ldots, M-1 \tag{A2}$$

(b) Assume the first alternative holds, and v satisfies (A1). Show that there exists a primal optimal solution $\tilde{x}$ with

$$\begin{aligned} a_1^T \tilde{x} &< b_1 \\ a_i^T \tilde{x} &\leq b_i, \qquad i = 2, \ldots, M \\ a_i^T \tilde{x} &= b_i, \qquad i = M+1, \ldots, N \\ a_i^T \tilde{x} &< b_i, \qquad i = N+1, \ldots, m. \end{aligned}$$

(c) Assume the second alternative holds, and w satisfies (A2). Show that there exists a dual optimal $\tilde{z}$ with

$$\begin{aligned} \tilde{z}_1 &> 0 \\ \tilde{z}_i &\geq 0, \qquad i = 2, \ldots, M \\ \tilde{z}_i &> 0, \qquad i = M+1, \ldots, N \\ \tilde{z}_i &= 0, \qquad i = N+1, \ldots, m. \end{aligned}$$

(d) Combine (b) and (c) to show that there exists a primal-dual optimal pair x, z, for which $b - Ax$ and z have at most $\tilde{M}$ common zeros, where $\tilde{M} < M$. If $\tilde{M} = 0$, x, z are strictly complementary and optimal, and we are done. Otherwise, we apply the argument given above, with x^*, z^* replaced by x, z, to show the existence of a strictly complementary pair of optimal solutions with less than $\tilde{M}$ common zeros in $b - Ax$ and z. Repeating the argument eventually gives a strictly complementary pair.

Goldman-Tucker Theorem for LO in standard form

Let the matrix $A : m \times n$ and the vectors $c \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$ be given. Apply the convex Farkas Lemma to prove the so-called Goldman-Tucker theorem for the LO problem:

$$\min\{c^T x : Ax = b,\ x \geq 0\}$$

when it admits an optimal solution. In other words, prove that there exists an optimal solution x^* and an optimal solution (y^*, s^*) of the dual LO problem

$$\max\{b^T y : A^T y + s = c,\ s \geq 0\}$$

such that

$$x^* + s^* > 0.$$

Proof. $Glodman - Tucker\ Theorem$ is equivalent to

$$\begin{array}{l} -t < 0 \\ -(x+s) + t1 \le 0 \\ c^T x \le b^T y \\ Ax = b \\ A^T y + s = c \\ x \in R^n_+,\ s \in R^n_+,\ y \in R^m,\ t \in R \end{array}$$

has a solution;
Using $Farkas\ Lemma$, we know this is equivalent to that there is no $\lambda_1 \in R^n_+,\ \lambda_2 \in R_+,\ \lambda_3 \in R^m,\ \lambda_4 \in R^n$, such that

$$\left\{ \begin{array}{c} -t + \lambda_1^T[-(x+s)+t1] + \lambda_2[c^T x - b^T y] - \lambda_3^T(Ax - b) - \lambda_4(A^T y + s - c) \ge 0 \\ \forall\ x \in R^n_+,\ s \in R^n_+,\ y \in R^m,\ t \in R \end{array} \right.$$

$$\Longrightarrow \left\{ \begin{array}{l} (-1 + \lambda_1^T 1)t + (-\lambda_1 + \lambda_2 c - A^T\lambda_3)^T x + (-\lambda_1 - \lambda_4)^T s + (-\lambda_2 b - A\lambda_4)^T y + (b^T\lambda_3 + c^T\lambda 4) \ge 0 \\ \forall\ x \in R^n_+,\ s \in R^n_+,\ y \in R^m,\ t \in R \end{array} \right.$$

$$\Longrightarrow \left\{ \begin{array}{l} -1 + \lambda_1^T 1 = 0 \\ -\lambda_1 + \lambda_2 c - A^T\lambda_3 \ge 0 \\ -\lambda_1 - \lambda_4 \ge 0 \\ -\lambda_2 b - A\lambda_4 = 0 \\ b^T\lambda_3 + c^T\lambda 4 \ge 0 \\ \lambda_1 \in R^n_+,\ \lambda_2 \in R_+,\ \lambda_3 \in R^m,\ \lambda_4 \in R^n \end{array} \right.$$

$-\lambda_1 - \lambda_4 \ge 0 \Longleftrightarrow \lambda_4 = -\lambda_1 - \lambda_5,\ \lambda_5 \in R^n_+$, so

$$\Longrightarrow \left\{ \begin{array}{l} -1 + \lambda_1^T 1 = 0 \\ -\lambda_1 + \lambda_2 c - A^T\lambda_3 \ge 0 \\ -\lambda_2 b + A\lambda_1 + A\lambda_5 = 0 \\ b^T\lambda_3 - c^T\lambda_1 - c^T\lambda_5 \ge 0 \\ \lambda_1 \in R^n_+,\ \lambda_2 \in R_+,\ \lambda_3 \in R^m,\ \lambda_5 \in R^n_+ \end{array} \right.$$

$$\Longrightarrow \left\{ \begin{array}{l} -1 + \lambda_1^T 1 = 0 \\ -\lambda_1 + \lambda_2 c - A^T\lambda_3 \ge 0 \\ A\lambda_1 + A\lambda_5 = \lambda_2 b \\ \lambda_2 b^T\lambda_3 - \lambda_2 c^T\lambda_1 - \lambda_2 c^T\lambda_5 \ge 0 \\ \lambda_1 \in R^n_+,\ \lambda_2 \in R_+,\ \lambda_3 \in R^m,\ \lambda_5 \in R^n_+ \end{array} \right.$$

$$\Longrightarrow \left\{ \begin{array}{l} -1 + \lambda_1^T 1 = 0 \\ -(\lambda_1 + \lambda_5)^T\lambda_1 + (\lambda_1^T + \lambda_5)\lambda_2 c - (\lambda_1 + \lambda_5)^T A^T\lambda_3 \ge 0 \\ \lambda_1^T A^T\lambda_3 + \lambda_5^T A^T\lambda_3 - \lambda_2 c^T\lambda_1 - \lambda_2 c^T\lambda_5 \ge 0 \\ \lambda_1 \in R^n_+,\ \lambda_2 \in R_+,\ \lambda_3 \in R^m,\ \lambda_5 \in R^n_+ \end{array} \right.$$

$$\Longrightarrow \left\{ \begin{array}{l} -1 + \lambda_1^T 1 = 0 \\ -(\lambda_1 + \lambda_5)^T\lambda_1 \ge 0 \\ \lambda_1 \in R^n_+,\ \lambda_2 \in R_+,\ \lambda_3 \in R^m,\ \lambda_5 \in R^n_+ \end{array} \right.$$

This is impossible. Thus proved the problem.

Que 3. Implement the simplex method using Bland's rule, and test the example on page 7-21. (The result should match these on pages 7-26 to 7-28)

Que 4. Consider the problem: $\min \phi(x) = x - \log(1+x)$. Note that the domain of ϕ is $(-1, \infty)$.

(a) Show that ϕ is strictly convex on its domain, and that $x = 0$ is the minimizer of ϕ.

(b) Show that the iterates from Newton's method satisfy the recursive relation

$$x_{k+1} = -x_k^2.$$

This is called quadratic convergence.

(c) For which starting points do the iterates of Newton's method converge to the minimizer?

Que 5. Implement either the infeasible primal-dual path-following algorithm (on pages 10-8 to 10-10) or the path following algorithm via self-dual embedding (on pages 11-23 to 11-25) and test the algorithm on randomly generated linear optimization problems.