EE364a Homework 5 solutions

4.15 *Relaxation of Boolean LP.* In a *Boolean linear program*, the variable x is constrained to have components equal to zero or one:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x_i \in \{0,1\}, \quad i=1,\ldots,n. \end{array} \qquad (1)$$

In general, such problems are very difficult to solve, even though the feasible set is finite (containing at most 2^n points).

In a general method called *relaxation*, the constraint that x_i be zero or one is replaced with the linear inequalities $0 \le x_i \le 1$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & 0 \le x_i \le 1, \quad i=1,\ldots,n. \end{array} \qquad (2)$$

We refer to this problem as the *LP relaxation* of the Boolean LP (4.67). The LP relaxation is far easier to solve than the original Boolean LP.

(a) Show that the optimal value of the LP relaxation (4.68) is a lower bound on the optimal value of the Boolean LP (4.67). What can you say about the Boolean LP if the LP relaxation is infeasible?

(b) It sometimes happens that the LP relaxation has a solution with $x_i \in \{0,1\}$. What can you say in this case?

Solution.

(a) The feasible set of the relaxation includes the feasible set of the Boolean LP. It follows that the Boolean LP is infeasible if the relaxation is infeasible, and that the optimal value of the relaxation is less than or equal to the optimal value of the Boolean LP.

(b) The optimal solution of the relaxation is also optimal for the Boolean LP.

4.60 *Log-optimal investment strategy.* We consider a portfolio problem with n assets held over N periods. At the beginning of each period, we re-invest our total wealth, redistributing it over the n assets using a fixed, constant, allocation strategy $x \in \mathbf{R}^n$, where $x \succeq 0$, $\mathbf{1}^T x = 1$. In other words, if $W(t-1)$ is our wealth at the beginning of period t, then during period t we invest $x_i W(t-1)$ in asset i. We denote by $\lambda(t)$ the total return during period t, *i.e.*, $\lambda(t) = W(t)/W(t-1)$. At the end of the N periods our wealth has been multiplied by the factor $\prod_{t=1}^{N} \lambda(t)$. We call

$$\frac{1}{N}\sum_{t=1}^{N} \log \lambda(t)$$

the *growth rate* of the investment over the N periods. We are interested in determining an allocation strategy x that maximizes growth of our total wealth for large N.

We use a discrete stochastic model to account for the uncertainty in the returns. We assume that during each period there are m possible scenarios, with probabilities π_j, $j = 1, \ldots, m$. In scenario j, the return for asset i over one period is given by p_{ij}. Therefore, the return $\lambda(t)$ of our portfolio during period t is a random variable, with m possible values $p_1^T x, \ldots, p_m^T x$, and distribution

$$\pi_j = \mathbf{prob}(\lambda(t) = p_j^T x), \quad j = 1, \ldots, m.$$

We assume the same scenarios for each period, with (identical) independent distributions. Using the law of large numbers, we have

$$\lim_{N\to\infty} \frac{1}{N} \log\left(\frac{W(N)}{W(0)}\right) = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} \log \lambda(t) = \mathbf{E} \log \lambda(t) = \sum_{j=1}^{m} \pi_j \log(p_j^T x).$$

In other words, with investment strategy x, the long term growth rate is given by

$$R_{\mathrm{lt}} = \sum_{j=1}^{m} \pi_j \log(p_j^T x).$$

The investment strategy x that maximizes this quantity is called the *log-optimal investment strategy*, and can be found by solving the optimization problem

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^{m} \pi_j \log(p_j^T x) \\ \text{subject to} & x \succeq 0, \quad \mathbf{1}^T x = 1, \end{array}$$

with variable $x \in \mathbf{R}^n$.

Show that this is a convex optimization problem.

Solution. Actually, there's not much to do in this problem. The constraints, $x \succeq 0$, $\mathbf{1}^T x = 1$, are clearly convex, so we just need to show that the objective is concave (since it is to be maximized). We can do that in just a few steps: First, note that log is concave, so $\log(p_j^T x)$ is concave in x (on the domain, which is the open halfspace $\{x \mid p_j^T x > 0\}$). Since $\pi_j \geq 0$, we conclude that the sum of concave functions

$$\sum_{j=1}^{m} \pi_j \log(p_j^T x)$$

is concave.

5.13 *Lagrangian relaxation of Boolean LP.* A *Boolean linear program* is an optimization problem of the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x_i \in \{0,1\}, \quad i = 1, \ldots, n, \end{array}$$

and is, in general, very difficult to solve. In exercise 4.15 we studied the LP relaxation of this problem,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & 0 \le x_i \le 1, \quad i = 1, \ldots, n, \end{array} \tag{3}$$

which is far easier to solve, and gives a lower bound on the optimal value of the Boolean LP. In this problem we derive another lower bound for the Boolean LP, and work out the relation between the two lower bounds.

(a) *Lagrangian relaxation.* The Boolean LP can be reformulated as the problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x_i(1 - x_i) = 0, \quad i = 1, \ldots, n, \end{array}$$

which has quadratic equality constraints. Find the Lagrange dual of this problem. The optimal value of the dual problem (which is convex) gives a lower bound on the optimal value of the Boolean LP. This method of finding a lower bound on the optimal value is called *Lagrangian relaxation.*

(b) Show that the lower bound obtained via Lagrangian relaxation, and via the LP relaxation (5.107), are the same. *Hint.* Derive the dual of the LP relaxation (5.107).

Solution.

(a) The Lagrangian is

$$\begin{aligned} L(x, \mu, \nu) &= c^T x + \mu^T (Ax - b) - \nu^T x + x^T \mathbf{diag}(\nu) x \\ &= x^T \mathbf{diag}(\nu) x + (c + A^T \mu - \nu)^T x - b^T \mu. \end{aligned}$$

Minimizing over x gives the dual function

$$g(\mu, \nu) = \begin{cases} -b^T \mu - (1/4) \sum_{i=1}^n (c_i + a_i^T \mu - \nu_i)^2 / \nu_i & \nu \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

where a_i is the ith column of A, and we adopt the convention that $a^2/0 = \infty$ if $a \neq 0$, and $a^2/0 = 0$ if $a = 0$.

The resulting dual problem is

$$\begin{array}{ll} \text{maximize} & -b^T\mu - (1/4)\sum_{i=1}^n (c_i + a_i^T\mu - \nu_i)^2/\nu_i \\ \text{subject to} & \nu \succeq 0, \quad \mu \succeq 0. \end{array}$$

In order to simplify this dual, we optimize analytically over ν, by noting that

$$\begin{aligned} \sup_{\nu_i \geq 0} \left(-\frac{(c_i + a_i^T\mu - \nu_i)^2}{\nu_i} \right) &= \begin{cases} 4(c_i + a_i^T\mu) & c_i + a_i^T\mu \leq 0 \\ 0 & c_i + a_i^T\mu \geq 0 \end{cases} \\ &= \min\{0, 4(c_i + a_i^T\mu)\}. \end{aligned}$$

This allows us to eliminate ν from the dual problem, and simplify it as

$$\begin{array}{ll} \text{maximize} & -b^T\mu + \sum_{i=1}^n \min\{0, c_i + a_i^T\mu\} \\ \text{subject to} & \mu \succeq 0. \end{array}$$

(b) We follow the hint. The Lagrangian and dual function of the LP relaxation are

$$\begin{aligned} L(x,u,v,w) &= c^Tx + u^T(Ax-b) - v^Tx + w^T(x - \mathbf{1}) \\ &= (c + A^Tu - v + w)^Tx - b^Tu - \mathbf{1}^Tw \\ g(u,v,w) &= \begin{cases} -b^Tu - \mathbf{1}^Tw & A^Tu - v + w + c = 0 \\ -\infty & \text{otherwise.} \end{cases} \end{aligned}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -b^Tu - \mathbf{1}^Tw \\ \text{subject to} & A^Tu - v + w + c = 0 \\ & u \succeq 0, v \succeq 0, w \succeq 0, \end{array}$$

which is equivalent to the Lagrange relaxation problem derived above. We conclude that the two relaxations give the same value.

6.2 *ℓ_1-, ℓ_2-, and ℓ_∞-norm approximation by a constant vector.* What is the solution of the norm approximation problem with one scalar variable $x \in \mathbf{R}$,

$$\text{minimize} \quad \|x\mathbf{1} - b\|,$$

for the ℓ_1-, ℓ_2-, and ℓ_∞-norms?

Solution.

(a) ℓ_2-norm: the average $\mathbf{1}^Tb/m$.

(b) ℓ_1-norm: the (or a) median of the coefficients of b.

(c) ℓ_∞-norm: the midrange point $(\max b_i - \min b_i)/2$.

Solutions to additional exercises

1. *Schur complements.* Consider a matrix $X = X^T \in \mathbf{R}^{n\times n}$ partitioned as

$$X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix},$$

where $A \in \mathbf{R}^{k\times k}$. If $\det A \neq 0$, the matrix $S = C - B^T A^{-1} B$ is called the *Schur complement* of A in X. Schur complements arise in many situations and appear in many important formulas and theorems. For example, we have $\det X = \det A \det S$. (You don't have to prove this.)

(a) The Schur complement arises when you minimize a quadratic form over some of the variables. Let $f(u, v) = [u^T \; v^T] X [u^T \; v^T]^T$, where $u \in \mathbf{R}^k$. Let $g(v)$ be the minimum value of f over u, *i.e.*, $g(v) = \inf_u f(u, v)$. Of course $g(v)$ can be $-\infty$. Show that if $A \succ 0$, we have $g(v) = v^T S v$.

(b) The Schur complement arises in several characterizations of positive definiteness or semidefiniteness of a block matrix. As examples we have the following three theorems:

- $X \succ 0$ if and only if $A \succ 0$ and $S \succ 0$.
- If $A \succ 0$, then $X \succeq 0$ if and only if $S \succeq 0$.
- $X \succeq 0$ if and only if $A \succeq 0$, $B^T(I - AA^\dagger) = 0$ and $C - B^T A^\dagger B \succeq 0$, where $A^\dagger$ is the pseudo-inverse of A. ($C - B^T A^\dagger B$ serves as a generalization of the Schur complement in the case where A is positive semidefinite but singular.)

Prove *one* of these theorems. (You can choose which one.)

(c) Recognizing Schur complements often helps to represent nonlinear convex constraints as linear matrix inequalities (LMIs). Consider the function

$$f(x) = (Ax + b)^T (P_0 + x_1 P_1 + \cdots + x_n P_n)^{-1} (Ax + b)$$

where $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, and $P_i = P_i^T \in \mathbf{R}^{m\times m}$, with domain

$$\mathbf{dom}\, f = \{x \in \mathbf{R}^n \mid P_0 + x_1 P_1 + \cdots + x_n P_n \succ 0\}.$$

This is the composition of the matrix fractional function and an affine mapping, and so is convex. Give an LMI representation of $\mathbf{epi}\, f$. That is, find a symmetric matrix $F(x, t)$, affine in (x, t), for which

$$x \in \mathbf{dom}\, f, \quad f(x) \leq t \iff F(x, t) \succeq 0.$$

Solution

(a) *If $A \succ 0$, then $g(v) = v^T S v$.*
We have $f(u,v) = u^T A u + 2v^T B u + v^T C v$. If $A \succ 0$, we can minimize f over u by setting the gradient with respect to u equal to zero. We obtain $u^\star(v) = -A^{-1}Bv$, and

$$g(v) = f(u^\star(v), v) = v^T(C - B^T A^{-1} B)v = v^T S v.$$

(b) *Positive definite and semidefinite block matrices.*

- *$X \succ 0$ if and only if $A \succ 0$ and $S \succ 0$.*
Suppose $X \succ 0$. Then $f(u,v) > 0$ for all non-zero (u,v), and in particular, $f(u,0) = u^T A u > 0$ for all non-zero u (hence, $A \succ 0$), and $f(-A^{-1}Bv, v) = v^T(C - B^T A^{-1} B)v > 0$ (hence, $S = C - B^T A^{-1} B \succ 0$). This proves the 'only if' part.
To prove the "if" part, we have to show that if $A \succ 0$ and $S \succ 0$, then $f(u,v) > 0$ for all nonzero (u,v) (that is, for all u, v that are not both zero). If $v \neq 0$, then it follows from (a) that

$$f(u,v) \geq \inf_u f(u,v) = v^T S v > 0.$$

If $v = 0$ and $u \neq 0$, $f(u,0) = u^T A u > 0$.
- *If $A \succ 0$, then $X \succeq 0$ if and only if $S \succeq 0$.*
From part (a) we know that if $A \succ 0$, then $\inf_u f(u,v) = v^T S v$. If $S \succeq 0$, then

$$f(u,v) \geq \inf_u f(u,v) = v^T S v \geq 0$$

for all u, v, and hence $X \succeq 0$. This proves the 'if'-part.
To prove the 'only if'-part we note that $f(u,v) \geq 0$ for all (u,v) implies that $\inf_u f(u,v) \geq 0$ for all v, *i.e.*, $S \succeq 0$.
- *$X \succeq 0$ if and only if $A \succeq 0$, $B^T(I - AA^\dagger) = 0$, $C - B^T A^\dagger B \succeq 0$.*
Suppose $A \in \mathbf{R}^{k \times k}$ with $\mathbf{rank}(A) = r$. Then there exist matrices $Q_1 \in \mathbf{R}^{k \times r}$, $Q_2 \in \mathbf{R}^{k \times (k-r)}$ and an invertible diagonal matrix $\Lambda \in \mathbf{R}^{r \times r}$ such that

$$A = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}^T,$$

and $[Q_1\ Q_2]^T [Q_1\ Q_2] = I$. The matrix

$$\begin{bmatrix} Q_1 & Q_2 & 0 \\ 0 & 0 & I \end{bmatrix} \in \mathbf{R}^{n \times n}$$

is nonsingular, and therefore

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0 \iff \begin{bmatrix} Q_1 & Q_2 & 0 \\ 0 & 0 & I \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 & 0 \\ 0 & 0 & I \end{bmatrix} \succeq 0$$

$$\iff \begin{bmatrix} \Lambda & 0 & Q_1^T B \\ 0 & 0 & Q_2^T B \\ B^T Q_1 & B^T Q_2 & C \end{bmatrix} \succeq 0$$

$$\iff Q_2^T B = 0, \begin{bmatrix} \Lambda & Q_1^T B \\ B^T Q_1 & C \end{bmatrix} \succeq 0.$$

We have $\Lambda \succ 0$ if and only if $A \succeq 0$. It can be verified that

$$A^\dagger = Q_1 \Lambda^{-1} Q_1^T, \quad I - AA^\dagger = Q_2 Q_2^T.$$

Therefore

$$Q_2^T B = 0 \iff Q_2 Q_2^T B = (I - A^\dagger A) B = 0.$$

Moreover, since Λ is invertible,

$$\begin{bmatrix} \Lambda & Q_1^T B \\ B^T Q_1 & C \end{bmatrix} \succeq 0 \iff \Lambda \succ 0, \;\; C - B^T Q_1 \Lambda^{-1} Q_1^T B = C - B^T A^\dagger B \succeq 0.$$

(c) The epigraph of f is the set of points (x, t) that satisfy $P_0 + x_1 P_1 + \cdots + x_n P_n \succ 0$ and

$$(Ax + b)^T (P_0 + x_1 P_1 + \cdots + x_n P_n)^{-1} (Ax + b) \le t.$$

Using the second result of part (b), we can write the second inequality as

$$\begin{bmatrix} t & (Ax+b)^T \\ (Ax+b) & P_0 + x_1 P_1 + \cdots + x_n P_n \end{bmatrix} \succeq 0.$$

This a linear matrix inequality in the variables x, t, *i.e.*, a convex constraint.

2. Formulate the following optimization problems as semidefinite programs. The variable is $x \in \mathbf{R}^n$; $F(x)$ is defined as

$$F(x) = F_0 + x_1 F_1 + x_2 F_2 + \cdots + x_n F_n$$

with $F_i \in \mathbf{S}^m$. The domain of f in each subproblem is $\mathbf{dom}\, f = \{x \in \mathbf{R}^n \mid F(x) \succ 0\}$.

(a) Minimize $f(x) = c^T F(x)^{-1} c$ where $c \in \mathbf{R}^m$.

(b) Minimize $f(x) = \max_{i=1,\ldots,K} c_i^T F(x)^{-1} c_i$ where $c_i \in \mathbf{R}^m$, $i = 1, \ldots, K$.

(c) Minimize $f(x) = \sup_{\|c\|_2 \le 1} c^T F(x)^{-1} c$.

(d) Minimize $f(x) = \mathbf{E}(c^T F(x)^{-1} c)$ where c is a random vector with mean $\mathbf{E}\, c = \bar{c}$ and covariance $\mathbf{E}(c - \bar{c})(c - \bar{c})^T = S$.

Solution.

(a)

$$
\begin{array}{ll}
\text{minimize} & t \\
\text{subject to} & \begin{bmatrix} F(x) & c \\ c^T & t \end{bmatrix} \succeq 0.
\end{array}
$$

(b)

$$
\begin{array}{ll}
\text{minimize} & t \\
\text{subject to} & \begin{bmatrix} F(x) & c_i \\ c_i^T & t \end{bmatrix} \succeq 0, \quad i = 1, \ldots, K.
\end{array}
$$

(c) $f(x) = \lambda_{\max}(F(x)^{-1})$, so $f(x) \leq t$ if and only if $F(x)^{-1} \preceq tI$. Using a Schur complement we get

$$
\begin{array}{ll}
\text{minimize} & t \\
\text{subject to} & \begin{bmatrix} F(x) & I \\ I & tI \end{bmatrix} \succeq 0.
\end{array}
$$

(d) $f(x) = \bar{c}^T F(x)^{-1} \bar{c} + \mathbf{tr}(F(x)^{-1} S)$. If we factor S as $S = \sum_{k=1}^m c_k c_k^T$ the problem is equivalent to

$$
\text{minimize} \quad \bar{c}^T F(x)^{-1} \bar{c} + \textstyle\sum_{k=1}^m c_k^T F(x)^{-1} c_k,
$$

which we can write as an SDP

$$
\begin{array}{ll}
\text{minimize} & t_0 + \sum_k t_k \\
\text{subject to} & \begin{bmatrix} F(x) & \bar{c} \\ \bar{c}^T & t_0 \end{bmatrix} \succeq 0 \\
& \begin{bmatrix} F(x) & c_k \\ c_k^T & t_k \end{bmatrix} \succeq 0, \quad k = 1, \ldots, m.
\end{array}
$$

3. *Optimality conditions and dual for log-optimal investment problem.*

(a) Show that the optimality conditions for the log-optimal investment problem described in exercise 4.60 can be expressed as: $\mathbf{1}^T x = 1$, $x \succeq 0$, and for each i,

$$
x_i > 0 \;\Rightarrow\; \sum_{j=1}^m \pi_j \frac{p_{ij}}{p_j^T x} = 1, \qquad x_i = 0 \;\Rightarrow\; \sum_{j=1}^m \pi_j \frac{p_{ij}}{p_j^T x} \leq 1.
$$

We can interpret this as follows. $p_{ij}/p_j^T x$ is a random variable, which gives the ratio of the investment gain with asset i only, to the investment gain with our mixed portfolio x. The optimality condition is that, for each asset we invest in, the expected value of this ratio is one, and for each asset we do not invest in, the expected value cannot exceed one. Very roughly speaking, this means our portfolio does as well as any of the assets that we choose to invest in, and cannot do worse than any assets that we do not invest in.

Hint. You can start from the simple criterion given in §4.2.3, or the KKT conditions, or additional exercise 1 from homework 4.

(b) In this part we will derive the dual of the log-optimal investment problem. We start by writing the problem as,

$$\begin{array}{ll} \text{minimize} & -\sum_{j=1}^m \pi_j \log y_j \\ \text{subject to} & y = P^T x, \quad x \succeq 0, \quad \mathbf{1}^T x = 1. \end{array}$$

Here, P has columns $p_1, \ldots, p_m$, and we have the introduced new variables $y_1, \ldots, y_m$, with the implicit constraint $y \succ 0$. We will associate dual variables ν, λ and ν_0 with the constraints $y = P^T x$, $x \succeq 0$, and $\mathbf{1}^T x = 1$, respectively. Defining $\tilde{\nu}_j = \nu_j/\nu_0$ for $j = 1, \ldots, m$, show that the dual problem can be written as

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^m \pi_j \log(\tilde{\nu}_j/\pi_j) \\ \text{subject to} & P\tilde{\nu} \preceq \mathbf{1}, \end{array}$$

with variable $\tilde{\nu}$. The objective here is the (negative) Kullback-Leibler divergence between the given distribution π and the dual variable $\tilde{\nu}$.

Solution.

(a) The problem is the same as minimizing $f(x) = -\sum_{j=1}^m \pi_j \log(p_j^T x)$ over the probability simplex. If x is feasible the optimality condition is that $\nabla f(x)^T (z - x) \geq 0$ for all z with $z \succeq 0$, $\mathbf{1}^T z = 1$. This holds if and only if for each k,

$$x_k > 0 \Rightarrow \frac{\partial f}{\partial x_k} = \min_{i=1,\ldots,n} \frac{\partial f}{\partial x_i} = \nabla f(x)^T x.$$

From $f(x) = -\sum_{j=1}^m \pi_j \log(p_j^T x)$, we get

$$\nabla f(x)_i = -\sum_{j=1}^m \pi_j (1/p_j^T x) p_{ij},$$

so

$$\nabla f(x)^T x = -\sum_{j=1}^m \pi_j (1/p_j^T x) p_j^T x = -1.$$

The optimality conditions are, $\mathbf{1}^T x = 1$, $x \succeq 0$, and for each i,

$$x_i > 0 \;\Rightarrow\; \sum_{j=1}^m \pi_j \frac{p_{ij}}{p_j^T x} = 1, \quad x_i = 0 \;\Rightarrow\; \sum_{j=1}^m \pi_j \frac{p_{ij}}{p_j^T x} \leq 1.$$

(b) The Lagrangian is

$$L(x, \nu, \lambda, \nu_0) = -\sum_{j=1}^m \pi_j \log y_j + \nu^T (y - P^T x) - \lambda^T x + \nu_0 (\mathbf{1}^T x - 1).$$

This is unbounded below unless $-\nu^T P^T - \lambda^T + \nu_0 \mathbf{1}^T = 0$. Since L is separable in y we can minimize it over y by minimizing over y_j. We find the minimum is obtained for $y_j = \pi_j/\nu_j$, so we have

$$g(\nu, \lambda, \nu_0) = 1 - \nu_0 + \sum_{j=1}^{m} \pi_j \log(\nu_j/\pi_j),$$

provided $P\nu \preceq \nu_0 \mathbf{1}$. Thus we can write the dual problem as

$$\begin{array}{ll} \text{maximize} & 1 - \nu_0 + \sum_{j=1}^m \pi_j \log(\nu_j/\pi_j) \\ \text{subject to} & P\nu \preceq \nu_0 \mathbf{1} \end{array}$$

with variables $\nu \in \mathbf{R}^m$, $\nu_0 \in \mathbf{R}$. This has implicit constraint $\nu \succ 0$.

We can further simplify this, by analytically optimizing over ν_0. From the constraint inequality we see that $\nu_0 > 0$. Defining $\tilde{\nu} = \nu/\nu_0$, we get the problem in variables $\tilde{\nu}$, ν_0

$$\begin{array}{ll} \text{maximize} & 1 - \nu_0 + \sum_{j=1}^m \pi_j \log(\tilde{\nu}_j \nu_0/\pi_j) \\ \text{subject to} & \nu_0 P\tilde{\nu} \preceq \nu_0 \mathbf{1}. \end{array}$$

Cancelling ν_0 from the constraint we get

$$\begin{array}{ll} \text{maximize} & 1 - \nu_0 + \log \nu_0 + \sum_{j=1}^m \pi_j \log(\tilde{\nu}_j/\pi_j) \\ \text{subject to} & P\tilde{\nu} \preceq \mathbf{1}. \end{array}$$

The optimal value of ν_0 is evidently $\nu_0 = 1$, so we end up with the dual problem

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^m \pi_j \log(\tilde{\nu}_j/\pi_j) \\ \text{subject to} & P\tilde{\nu} \preceq \mathbf{1}. \end{array}$$

4. *Log-optimal investment strategy.* In this problem you will solve a specific instance of the log-optimal investment problem described in exercise 4.60, with $n = 5$ assets and $m = 10$ possible outcomes in each period. The problem data are defined in `log_opt_invest.m`, with the rows of the matrix `P` giving the asset return vectors p_j^T. The outcomes are equiprobable, *i.e.*, we have $\pi_j = 1/m$. Each column of the matrix `P` gives the return of the associated asset in the different posible outcomes. You can examine the columns to get an idea of the types of assets. For example, the last asset gives a fixed and certain return of 1%; the first asset is a very risky one, with occasional large return, and (more often) substantial loss.

   Find the log-optimal investment strategy $x^\star$, and its associated long term growth rate $R_{\rm lt}^\star$. Compare this to the long term growth rate obtained with a uniform allocation strategy, *i.e.*, $x = (1/n)\mathbf{1}$, and also with a pure investment in each asset.

   For the optimal investment strategy, and also the uniform investment strategy, plot 10 sample trajectories of the accumulated wealth, *i.e.*, $W(T) = W(0)\prod_{t=1}^{T} \lambda(t)$, for $T = 0, \ldots, 200$, with initial wealth $W(0) = 1$.

To save you the trouble of figuring out how to simulate the wealth trajectories or plot them nicely, we've included the simulation and plotting code in `log_opt_invest.m`; you just have to add the code needed to find $x^\star$.

Hint: The current version of `cvx` doesn't handle the logarithm, but you can use `geomean()` to solve the problem.

Solution.

(a) The following code was used to solve this problem:

```
P = [3.5000    1.1100    1.1100    1.0400    1.0100;
     0.5000    0.9700    0.9800    1.0500    1.0100;
     0.5000    0.9900    0.9900    0.9900    1.0100;
     0.5000    1.0500    1.0600    0.9900    1.0100;
     0.5000    1.1600    0.9900    1.0700    1.0100;
     0.5000    0.9900    0.9900    1.0600    1.0100;
     0.5000    0.9200    1.0800    0.9900    1.0100;
     0.5000    1.1300    1.1000    0.9900    1.0100;
     0.5000    0.9300    0.9500    1.0400    1.0100;
     3.5000    0.9900    0.9700    0.9800    1.0100];

[m,n] = size(P);

% Find log-optimal investment policy
cvx_begin
   variable x_opt(n)
   maximize(geomean(P*x_opt))
   sum(x_opt) == 1
   x_opt >= 0
cvx_end

x_opt
x_unif = ones(n,1)/n;
R_opt = sum(log(P*x_opt))/m
R_unif = sum(log(P*x_unif))/m
```

It was found that the log-optimal investment strategy is:

$$x_{\mathrm{opt}} = (0.0580,\ 0.4000,\ 0.2923,\ 0.2497,\ 0.0000)\,.$$

This strategy achieves a long term growth rate $R^*_{\mathrm{lt}} = 0.0231$. In contrast, the uniform allocation strategy achieves a growth rate of $R_{\mathrm{unif}} = 0.0114$.

Clearly asset 1 is a high-risk asset. The amount that we invest in this asset will grow by a factor of 3.50 with probability 20% and will be halved with probability

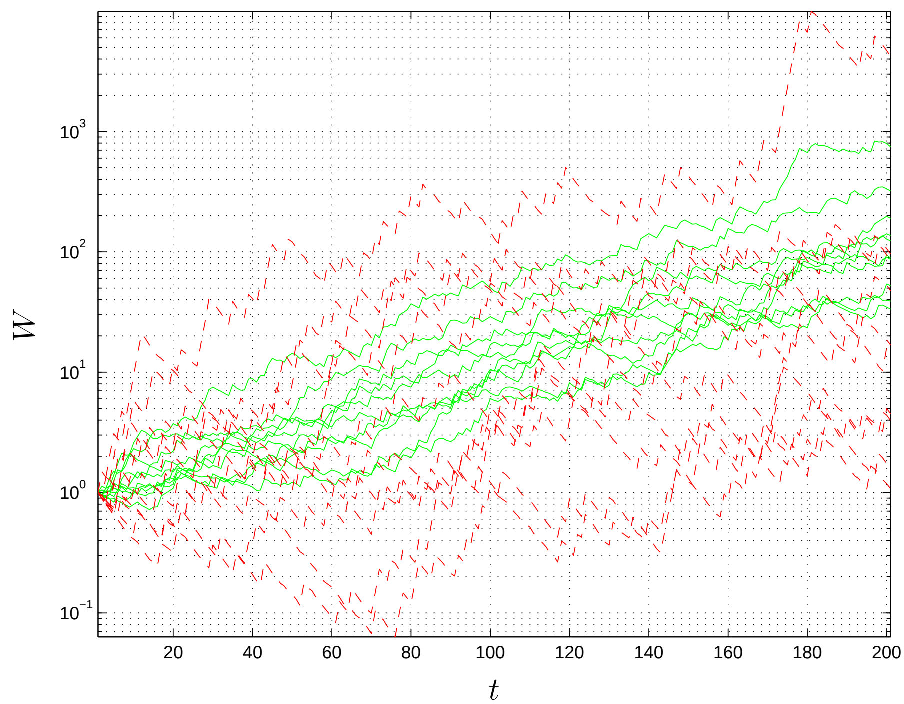
80%. On the other hand, asset 5 is an asset with a certain return of 1% per time period. Finally, assets 2, 3 and 4 are low-risk assets. It turns out that the log-optimal policy in this case is to invest very little wealth in the high-risk asset and no wealth on the sure asset and to invest most of the wealth in asset 2.

(b) The following code was used to generate the random event sequences and the trajectory plots:

```
% Generate random event sequences
rand('state',0);
N = 10; T = 200;
w_opt = []; w_unif = [];
for i = 1:N
   events = ceil(rand(1,T)*m);
   P_event = P(events,:);
   w_opt = [w_opt [1; cumprod(P_event*x_opt)]];
   w_unif = [w_unif [1; cumprod(P_event*x_unif)]];
end

% Plot wealth versus time
figure
semilogy(w_opt,'g')
hold on
semilogy(w_unif,'r--')
grid
axis tight
xlabel('time')
ylabel('wealth')
```

This generates the following plot:



The log-optimal investment policy consistently increases the wealth. On the other hand the uniform allocation policy generates quite random trajectories, a few with very large increases in wealth, and many with poor performance. This is due to the fact that with this policy 20% of the wealth is invested in the high-risk asset.

5. *Maximizing house profit in a gamble and imputed probabilities.* A set of n participants bet on which one of m outcomes, labeled $1, \ldots, m$, will occur. Participant i offers to purchase up to $q_i > 0$ gambling contracts, at price $p_i > 0$, that the true outcome will be in the set $S_i \subset \{1, \ldots, m\}$. The house then sells her x_i contracts, with $0 \leq x_i \leq q_i$. If the true outcome j is in S_i, then participant i receives \$1 per contract, *i.e.*, x_i. Otherwise, she loses, and receives nothing. The house collects a total of $x_1 p_1 + \cdots + x_n p_n$, and pays out an amount that depends on the outcome j,

$$\sum_{j \in S_i} x_i.$$

The difference is the house profit.

   (a) *Optimal house strategy.* How should the house decide on x so that its worst-case profit (over the possible outcomes) is maximized? (The house determines x after examining all the participant offers.)

   (b) *Imputed probabilities.* Suppose $x^\star$ maximizes the worst-case house profit. Show that there exists a probability distribution π on the possible outcomes (*i.e.*, $\pi \in \mathbf{R}_+^m$, $\mathbf{1}^T \pi = 1$) for which $x^\star$ also maximizes the expected house profit. Explain how to find π.

Hint. Formulate the problem in part (a) as an LP; you can construct π from optimal dual variables for this LP.

Remark. Given π, the 'fair' price for offer i is $p_i^{\text{fair}} = \sum_{j \in S_i} \pi_j$. All offers with $p_i > p_i^{\text{fair}}$ will be completely filled (*i.e.*, $x_i = q_i$); all offers with $p_i < p_i^{\text{fair}}$ will be rejected (*i.e.*, $x_i = 0$).

Remark. This exercise shows how the probabilities of outcomes (*e.g.*, elections) can be guessed from the offers of a set of gamblers.

(c) *Numerical example.* Carry out your method on the simple example below with $n = 5$ participants, $m = 5$ possible outcomes, and participant offers

Participant i	p_i	q_i	S_i
1	0.50	10	{1,2}
2	0.60	5	{4}
3	0.60	5	{1,4,5}
4	0.60	20	{2,5}
5	0.20	10	{3}

Compare the optimal worst-case house profit with the worst-case house profit, if all offers were accepted (*i.e.*, $x_i = q_i$). Find the imputed probabilities.

Solution.

(a) The worst-case house profit is

$$p^T x - \max_{j=1,\dots,m} \sum_{j \in S_i} x_i,$$

which is a piecewise-linear concave function of x. To find the x that maximizes the worst-case profit, we solve the problem,

$$\begin{array}{ll} \text{maximize} & p^T x - \max_{j=1,\dots,m} a_j^T x \\ \text{subject to} & 0 \preceq x \preceq q, \end{array}$$

with variable x. a_j^T are the rows of the subset matrix A, with

$$A_{ji} = \begin{cases} 1 & j \in S_i \\ 0 & \text{otherwise.} \end{cases}$$

(b) The problem from part (a) can be expressed as

$$\begin{array}{ll} \text{maximize} & p^T x - t \\ \text{subject to} & t\mathbf{1} \succeq Ax \\ & 0 \preceq x \preceq q, \end{array} \qquad (4)$$

where t is a new scalar variable. The Lagrangian is

$$L(x, t, \lambda_1, \lambda_2, \lambda_3) = t - p^T x + \lambda_1^T (Ax - t\mathbf{1}) - \lambda_2^T x + \lambda_3^T (x - q).$$

This is bounded below if and only if $\mathbf{1}^T\lambda_1 = 1$, and $A^T\lambda_1 - \lambda_2 + \lambda_3 = p$. The dual can be written as

$$\begin{array}{ll} \text{maximize} & -q^T\lambda_3 \\ \text{subject to} & \mathbf{1}^T\lambda_1 = 1 \\ & A^T\lambda_1 - \lambda_2 + \lambda_3 = p \\ & \lambda_1 \succeq 0, \quad \lambda_2 \succeq 0, \quad \lambda_3 \succeq 0, \end{array} \tag{5}$$

with variables λ_1, λ_2, and λ_3. Notice that λ_1 must satisfy $\mathbf{1}^T\lambda_1 = 1$, and $\lambda_1 \succeq 0$, hence it is a probability distribution.

Suppose $x^\star_{\text{wc}}$, $t^\star$, $\lambda_1^\star$, $\lambda_2^\star$, and $\lambda_3^\star$ are primal and dual optimal for problem (4), and let us set $\pi = \lambda_1^\star$. To maximize the expected house profit we solve the problem,

$$\begin{array}{ll} \text{maximize} & p^Tx - \pi^TAx \\ \text{subject to} & 0 \preceq x \preceq q. \end{array} \tag{6}$$

Let $b_1^T, \ldots, b_n^T$ be the rows of A^T. We know that a point $x^\star_{\text{e}}$ is optimal for problem (6) if and only if $x^\star_{\text{e}i} = q_i$ when $p_i - b_i^T\pi > 0$, $x^\star_{\text{e}i} = 0$ when $p_i - b_i^T\pi < 0$, and $0 \leq x^\star_{\text{e}i} \leq q_i$ when $p_i - b_i^T\pi = 0$.

To see why $x^\star_{\text{e}} = x^\star_{\text{wc}}$, let us take a look at one of the KKT conditions for problem (4). This can be written as

$$p - A^T\pi = \lambda_3^\star - \lambda_2^\star,$$

with $\pi = \lambda_1^\star$. If $p_i - b_i^T\pi > 0$, then we must have $\lambda^\star_{3i} - \lambda^\star_{2i} > 0$, which means that $\lambda^\star_{2i} = 0$ and $\lambda^\star_{3i} > 0$ (by complementary slackness), and so $x^\star_{\text{wc}i} = q_i$. Similarly, if $p_i - b_i^T\pi < 0$, then $\lambda^\star_{3i} - \lambda^\star_{2i} < 0$, which means that $\lambda^\star_{2i} > 0$ and $\lambda^\star_{3i} = 0$, and so $x^\star_{\text{wc}i} = 0$. Finally, when $p_i - b_i^T\pi = 0$, we must have $\lambda^\star_{2i} = 0$ and $\lambda^\star_{3i} = 0$, and so $0 \leq x^\star_{\text{wc}i} \leq q_i$.

In summary, in order to find a probability distribution on the possible outcomes for which the same $x^\star$ maximizes both the worst-case as well as the expected house profit, we solve the dual LP (5), and set $\pi = \lambda_1^\star$.

(c) The following `cvx` code solves the problem.

```
% solution for gambling problem
A = [1 0 1 0 0;
     1 0 0 1 0;
     0 0 0 0 1;
     0 1 1 0 0;
     0 0 1 1 0];

p = [0.5; 0.6; 0.6; 0.6; 0.2];
q = [10; 5; 5; 20; 10];
n = 5; m = 5;
```

```
cvx_begin
    variables x(n) t
    dual variable lambda1
    maximize (p'*x-t)
    subject to
        lambda1: A*x <= t
        x >= 0
        x <= q
cvx_end

% optimal worst case house profit
pwc = cvx_optval
% optimal worst case profit if all offer are accepted
pwc_accept = p'*q-max(A*q)
% imputed probabilities
pi = lambda1
% fair prices
pfair = A'*pi
% optimal purchase quantities
xopt = x
```

Our results are summarized in the following table. We find that the optimal worst case house profit is 3.5, and the worst case house profit, if all offers are accepted, is -5. The imputed probabilities are

$$\pi = (0.1145,\ 0.3855,\ 0.0945,\ 0.1910,\ 0.2145)\,.$$

The associated fair prices and optimal contract numbers are shown below.

Participant i	p_i	p_i^{fair}	q_i	x_i	S_i
1	0.50	0.5000	10	5	{1,2}
2	0.60	0.1910	5	5	{4}
3	0.60	0.5200	5	5	{1,4,5}
4	0.60	0.6000	20	5	{2,5}
5	0.20	0.0945	10	10	{3}

6. *Heuristic suboptimal solution for Boolean LP.* This exercise builds on exercises 4.15 and 5.13, which involve the Boolean LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x_i \in \{0,1\}, \quad i = 1,\ldots,n, \end{array}$$

with optimal value $p^\star$. Let $x^{\rm rlx}$ be a solution of the LP relaxation

$$\begin{array}{ll} \mbox{minimize} & c^T x \\ \mbox{subject to} & Ax \preceq b \\ & 0 \preceq x \preceq \mathbf{1}, \end{array}$$

so $L = c^T x^{\rm rlx}$ is a lower bound on $p^\star$. The relaxed solution $x^{\rm rlx}$ can also be used to guess a Boolean point $\hat{x}$, by rounding its entries, based on a threshold $t \in [0,1]$:

$$\hat{x}_i = \left\{ \begin{array}{ll} 1 & x_i^{\rm rlx} \geq t \\ 0 & \mbox{otherwise}, \end{array} \right.$$

for $i = 1, \ldots, n$. Evidently $\hat{x}$ is Boolean (*i.e.*, has entries in $\{0,1\}$). If it is feasible for the Boolean LP, *i.e.*, if $A\hat{x} \preceq b$, then it can be considered a guess at a good, if not optimal, point for the Boolean LP. Its objective value, $U = c^T \hat{x}$, is an upper bound on $p^\star$. If U and L are close, then $\hat{x}$ is nearly optimal; specifically, $\hat{x}$ cannot be more than $(U - L)$-suboptimal for the Boolean LP.

This rounding need not work; indeed, it can happen that for all threshold values, $\hat{x}$ is infeasible. But for some problem instances, it can work well.

Of course, there are many variations on this simple scheme for (possibly) constructing a feasible, good point from $x^{\rm rlx}$.

Finally, we get to the problem. Generate problem data using

```
rand('state',0);
n=100;
m=300;
A=rand(m,n);
b=A*ones(n,1)/2;
c=-rand(n,1);
```

You can think of x_i as a job we either accept or decline, and $-c_i$ as the (positive) revenue we generate if we accept job i. We can think of $Ax \preceq b$ as a set of limits on m resources. A_{ij}, which is positive, is the amount of resource i consumed if we accept job j; b_i, which is positive, is the amount of resource i available.

Find a solution of the relaxed LP and examine its entries. Note the associated lower bound L. Carry out threshold rounding for (say) 100 values of t, uniformly spaced over $[0,1]$. For each value of t, note the objective value $c^T\hat{x}$ and the maximum constraint violation $\max_i (A\hat{x} - b)_i$. Plot the objective value and the maximum violation versus t. Be sure to indicate on the plot the values of t for which $\hat{x}$ is feasible, and those for which it is not.

Find a value of t for which $\hat{x}$ is feasible, and gives minimum objective value, and note the associated upper bound U. Give the gap $U - L$ between the upper bound on $p^\star$

and the lower bound on $p^\star$. If you define vectors `obj` and `maxviol`, you can find the upper bound as `U=min(obj(find(maxviol<=0)))`.

Solution.

The following Matlab code finds the solution

```
% generate data for boolean LP relaxation & heuristic
rand('state',0);
n=100;
m=300;
A=rand(m,n);
b=A*ones(n,1)/2;
c=-rand(n,1);

% solve LP relaxation
cvx_begin
   variable x(n)
   minimize (c'*x)
   subject to
       A*x <= b
       x>=0
       x<=1
cvx_end
xrlx = x;
L=cvx_optval;

% sweep over threshold & round
thres=0:0.01:1;
maxviol = zeros(length(thres),1);
obj = zeros(length(thres),1);
for i=1:length(thres)
   xhat = (xrlx>=thres(i));
   maxviol(i) = max(A*xhat-b);
   obj(i) = c'*xhat;
end

% find least upper bound and associated threshold
i_feas=find(maxviol<=0);
U=min(obj(i_feas))
t=min(i_feas);
min_thresh=thres(t)

% plot objective and max violation versus threshold
```

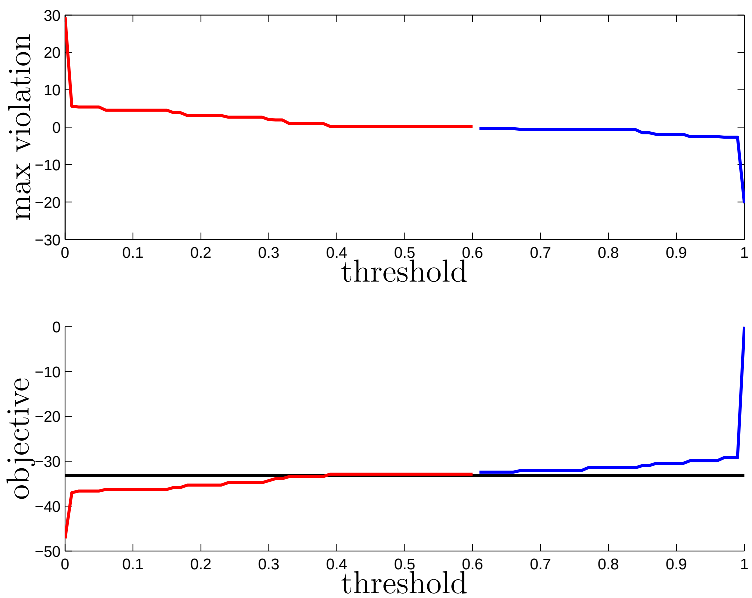


Figure 1 Plots of violation and objective vs threshold rule.

```
subplot(2,1,1)
plot(thres(1:t-1),maxviol(1:t-1),'r',thres(t:end),maxviol(t:end),'b','linewidth',2)
xlabel('threshold');
ylabel('max violation');
subplot(2,1,2)
hold on; plot(thres,L*ones(size(thres)),'k','linewidth',2);
plot(thres(1:t-1),obj(1:t-1),'r',thres(t:end),obj(t:end),'b','linewidth',2);
xlabel('threshold');
ylabel('objective');
```

The lower bound found from the relaxed LP is $L = -33.1672$. We find that the threshold value $t = 0.6006$ gives the best (smallest) objective value for feasible $\hat{x}$: $U = -32.4450$. The difference is 0.7222. So $\hat{x}$, with $t = 0.6006$, can be no more than 0.7222 suboptimal.

In figure 1, the red lines indicate values for thresholding values which give infeasible $\hat{x}$, and the blue lines correspond to feasible $\hat{x}$. We see that the maximum violation decreases as the threshold is increased. This occurs because the constraint matrix A only has nonnegative entries. At a threshold of 0, all jobs are selected, which is an infeasible solution. As we increase the threshold, projects are removed in sequence (without adding new projects), which monotonically decreases the maximum violation. For a general boolean LP, the corresponding plots need not exhibit monotonic behavior.