Convex Optimization

June 11, 2025

Contents

0.1 Convex Optimization Exercises

Problem:
4.1 Consider the optimization problem

$$\begin{array}{rl} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & 2x_1 + x_2 \geq 1 \\ & x_1 + 3x_2 \geq 1 \\ & x_1 \geq 0, \quad x_2 \geq 0. \end{array}$$

Make a sketch of the feasible set. For each of the following objective functions, give the optimal set and the optimal value.

(a) $f_0(x_1, x_2) = x_1 + x_2$.

(b) $f_0(x_1, x_2) = -x_1 - x_2$.

(c) $f_0(x_1, x_2) = x_1$.

(d) $f_0(x_1, x_2) = \max\{x_1, x_2\}$.

(e) $f_0(x_1, x_2) = x_1^2 + 9x_2^2$.

Solution:

(a) The optimal set consists of one point $(\frac{2}{5}, \frac{1}{5})$ and the optimal value is $\frac{3}{5}$. Consider the first intersection of a growing $L1$ square.

(b) The optimal set is empty and the optimal value is unbounded below.

(c) The optimal set is the set $\{(x, y)|x = 0, y \geq 0\}$, and the optimal value is 0.

(d) The optimal set is one point $(\frac{1}{3}, \frac{1}{3})$ and the optimal value is $\frac{1}{3}$. Consider the first intersection of a growing square axis aligned.

(e) The optimal set is one point $(\frac{1}{2}, \frac{1}{6})$, with optimal value: $\frac{1}{2}$. Consider a growing ellipse $(\frac{x_1}{r})^2 + (\frac{x_2}{r/3})^2$ which helps us to identify the first intersection should lie on the flatter line. More precise we can say $\nabla f_0(x^*) = (1, 3)$ is perpendicular to the line, implying the level set is tangent to the boundary.

Above we use growing level sets to intuitively identify minimal points.

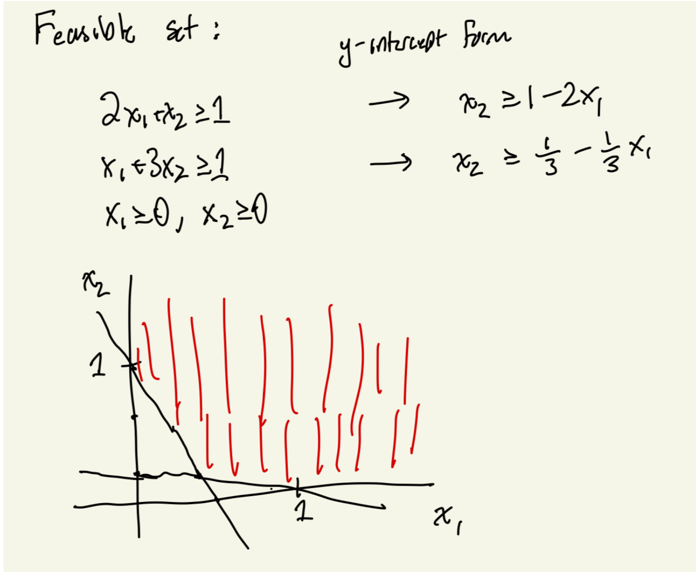

Figure 1: Enter Caption

Problem: 4.2 Consider the optimization problem

$$\text{minimize} \quad f_0(x) = -\sum_{i=1}^{m} \log\left(b_i - a_i^T x\right)$$

with domain dom $f_0 = \{x \mid Ax \prec b\}$, where $A \in \mathbb{R}^{m\times n}$ (with rows a_i^T). We assume that dom f_0 is nonempty. Prove the following facts (which include the results quoted without proof on page **141**).

(a) dom f_0 is unbounded if and only if there exists a $v \neq 0$ with $Av \preceq 0$

(b) f_0 is unbounded below if and only if there exists a v with $Av \preceq 0$, $Av \neq 0$. *Hint:* There exists a v such that $Av \preceq 0$, $Av \neq 0$ if and only if there exists no $z \succ 0$ such that $A^T z = 0$. This follows from the theorem of alternatives in example **2.21**, page **50**.

(c) If f_0 is bounded below then its minimum is attained, i.e., there exists an x that satisfies the optimality condition (4.23).

(d) The optimal set is affine: $X_{\text{opt}} = \{x^\star + v \mid Av = 0\}$, where $x^\star$ is any optimal point.

Solution:

(a) (**Note:** good analysis exercise)

(b) TODO.

(a) $\Rightarrow$ Suppose for the sake of contradiction that all $v \neq 0$, we have $Av > 0$ and $\mathbf{dom} f_0$ is unbounded. Let us consider a sequence of elements in the domain $\{x_k\}$ such that $\|x_k\| \to \infty$. Let us consider $v_k = \frac{x_k}{\|x_k\|}$, and we have $\|v_k\| = 1$. By BW, we have there is a convergent subsequence $\{v_{a_i}\} \to v^\star$.

From assumption, $Av^\star \succ 0$ and $(Av^\star)_{r'} = z^* > 0$. For $\epsilon = \frac{z^\star}{\|A\|}$, there exists k such that for all $i > k$, we have that $\|v^\star - v_{a_i}\| < \epsilon$.

$$\begin{aligned}
(Av_{a_i})_{r'} &= (Av^\star - A(v^\star - v_{a_i})) \\
&\geq z^\star - \|A\|\|v^\star - v_{a_i}\| \\
&\geq \frac{1}{2} z^\star
\end{aligned}$$

Note that there is some B such that $\frac{Bz^\star}{2} > b_{r'}$. Note that in the sequence $\{v_{a_i}\}$ there is some k' such that $i > k'$, $\|x_{a_i}\| \geq B$. Consider $k \leftarrow \min(k, k')$, then for all $i > k$ we have that $Ax_{a_i} \preccurlyeq b$.

$\Leftarrow$ Suppose there exists a $v \neq 0$ with $Av \preccurlyeq 0$. Then we can consider $x \in \mathbf{dom} f_0$ since the domain is non empty, and $x + zv, z \in \mathbb{Z}$. Clearly $x + zv$ is unbounded.

(b) $\Rightarrow$ Suppose f_0 is unbounded below. Then we have some sequence $\{x_n\}, \|f_0(x_n)\| \to \infty$.

$\Leftarrow$ Suppose there exists such a v with $Av \preceq 0$ and $Av \neq 0$. Then for any solution x, we can make $\|f_0(x+vt)\| \to -\infty$.

(c)

(d) Let S be an optimal set. Clearly $X_{opt} \subset S$. Suppose there is some optimal point $x^\star + v'$.

Problem: 4.3 Prove that $x^\star = (1, 1/2, -1)$ is optimal for the optimization problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2} x^T P x + q^T x + r \\ \text{subject to} & -1 \leq x_i \leq 1, \quad i = 1, 2, 3, \end{array}$$

where

$$P = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}, \quad q = \begin{bmatrix} -22.0 \\ -14.5 \\ 13.0 \end{bmatrix}, \quad r = 1.$$

Solution:

Note that our objective f_0 is differentiable. It suffices to show from section "Optimality criterion for differentiable f_0", that $\nabla f^T (y - x) \geq 0$ for all $y \in X$ with x as our optimal point. Evaluating this we have:

$$\begin{pmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{pmatrix} \begin{pmatrix} 1 \\ 1/2 \\ -1 \end{pmatrix} + \begin{pmatrix} -22 \\ -14.5 \\ 13.0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$$

Now we have to show that :

$$\begin{aligned} \begin{pmatrix} -1 & 0 & 2 \end{pmatrix} (y - x) &= -(y_1 - 1) + 2(y_3 + 1) \\ &= -y_1 + 2y_3 + 3 \\ -y_1 \geq -1, 2y_3 &\geq -2 \\ &\geq 0 \end{aligned}$$

Thus we are done. Geometrically this is a hyperplane supporting the entire feasible set at the optimality point.

Problem: 4.4 [P. Parrilo] *Symmetries and convex optimization.* Suppose $\mathcal{G} = \{Q_1, \dots, Q_k\} \subset \mathbb{R}^{n \times n}$ is a group, i.e., closed under products and inverse. We say that the function $f : \mathbb{R}^n \to \mathbb{R}$ is $\mathcal{G}$-*invariant*, or *symmetric* with respect to $\mathcal{G}$, if $f(Q_i x) = f(x)$ holds for all x and $i = 1, \dots, k$.

We define $\overline{x} = \frac{1}{k}\sum_{i=1}^{k} Q_i x$, which is the average of x over its $\mathcal{G}$-orbit. We define the *fixed subspace* of $\mathcal{G}$ as

$$\mathcal{F} = \{x \mid Q_i x = x,\ i = 1, \dots, k\}.$$

(a) Show that for any $x \in \mathbb{R}^n$, we have $\overline{x} \in \mathcal{F}$.

(b) Show that if $f : \mathbb{R}^n \to \mathbb{R}$ is convex and $\mathcal{G}$-invariant, then $f(\overline{x}) \le f(x)$.

(c) We say the optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le 0, \quad i = 1, \dots, m \end{array}$$

is $\mathcal{G}$-invariant if the objective f_0 is $\mathcal{G}$-invariant, and the feasible set is $\mathcal{G}$-invariant, which means

$$f_1(x) \le 0, \dots, f_m(x) \le 0 \quad \Longrightarrow \quad f_1(Q_i x) \le 0, \dots, f_m(Q_i x) \le 0,$$

for $i = 1, \dots, k$. Show that if the problem is convex and $\mathcal{G}$-invariant, and there exists an optimal point, then there exists an optimal point in $\mathcal{F}$. In other words, we can adjoin the equality constraints $x \in \mathcal{F}$ to the problem, without loss of generality.

(d) As an example, suppose f is convex and symmetric, i.e., $f(Px) = f(x)$ for every permutation P. Show that if f has a minimizer, then it has a minimizer of the form $\alpha\mathbf{1}$. (This means to minimize f over $x \in \mathbb{R}^n$, we can just as well minimize $f(\alpha\mathbf{1})$ over $t \in \mathbb{R}$.)

Solution: Notes:

- Standard
- Standard
- Standard
- Interesting insight/argument

(a) We claim that for any $j \in [k]$, $\{Q_j Q_i | i = 1, ..., k\}$ is the same set as $\{Q_i | i = 1, ..., k\}$. Suppose $Q_j Q_{i_1} = Q_j Q_{i_2}$ and $Q_{i_1} \ne Q_{i_2}$. Then since Q_j is invertible we have that $Q_{i_1} = Q_{i_2}$. Therefore $Q_j Q_i$ map to different Q from the original set. But both sets have precisely the same elements, therefore both sets must be identical.

So, $\overline{x} = \frac{1}{k}\sum_{i=1}^{k} Q_i x = \frac{1}{k}\sum_{i=1}^{k} Q_j Q_i x = Q_j \overline{x}$.

(b) We have that

$$\begin{aligned} f(\overline{x}) &= f(\frac{1}{k}\sum_{i=1}^{k} Q_i x) \\ &\le \frac{1}{k} f(Q_1 x) + ... + \frac{1}{k} f(Q_k x) && \text{by convexity} \\ &= \frac{1}{k} f(x) + ... + \frac{1}{k} f(x) && \text{by } \mathcal{G}\text{-invariance} \\ &= f(x) \end{aligned}$$

(c) Let $x^\star$ be an optimal point that is not contained in $\mathcal{F}$. Then $f(\overline{x^\star}) \le f(x)$, which implies that $\overline{x^\star}$ is an optimal point. But recall that for any $x \in \mathbb{R}^n$, we have that $Q_j \overline{x} = \overline{x}$ for all $j \in [k]$. Therefore $\overline{x^*} \in \mathcal{F}$. We can freely adjoin the constraints of membership to $\mathcal{F}$ to the problem.

(d) First we have that f is $P-$invariant. Let $x^\star$ be an optimum point. Then $\overline{x^\star}$ is also an optimum point. We have that $P\overline{x^\star} = \overline{x^\star}$ for all $P \in S_n$. But this means that $\overline{x^\star} = \alpha\mathbf{1}$ since the vector is invariant to all permutations.

Problem: 4.5 Show that the following three convex problems are equivalent. Carefully explain how the solution of each problem is obtained from the solution of the other problems. The problem data are the matrix $A \in \mathbb{R}^{m\times n}$ (with rows a_i^T), the vector $b \in \mathbb{R}^m$, and the constant $M > 0$.

(a) **The robust least-squares problem**

$$\text{minimize} \quad \sum_{i=1}^{m} \phi(a_i^T x - b_i),$$

with variable $x \in \mathbb{R}^n$, where $\phi : \mathbb{R} \to \mathbb{R}$ is defined as

$$\phi(u) = \begin{cases} u^2 & |u| \le M \\ M(2|u| - M) & |u| > M \end{cases}$$

(This function is known as the *Huber penalty function*; see §6.1.2.)

(b) **The least-squares problem with variable weights**

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{i=1}^{m} \frac{(a_i^T x - b_i)^2}{w_i + 1} + M^2 \mathbf{1}^T w \\ \text{subject to} & w \succeq 0, \end{array}$$

with variables $x \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$, and domain $\mathcal{D} = \{(x, w) \in \mathbb{R}^n \times \mathbb{R}^m \mid w \succ -1\}$.

Hint. Optimize over w assuming x is fixed, to establish a relation with the problem in part (a).
(This problem can be interpreted as a weighted least-squares problem in which we are allowed to adjust the weight of the ith residual. The weight is one if $w_i = 0$, and decreases if we increase w_i. The second term in the objective penalizes large values of w, i.e., large adjustments of the weights.)

(c) **The quadratic program**

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{i=1}^{m} (u_i^2 + 2Mv_i) \\ \text{subject to} & -u = Ax - b \preceq u + v \\ & 0 \preceq u \preceq M\mathbf{1} \\ & v \succeq 0. \end{array}$$

Notes: Good exercise on decomposing optimization problems. **Solution:**

(a) $a \Leftrightarrow b$:

Let us consider an optimal solution to $b : (x^\star, w^\star)$. Let us consider the summation:

$$\sum_{i=1}^{m} \frac{(a_i^T x - b_i)^2}{w_i + 1} + M^2 w_i$$

For the particular x^* at i we can calculate how much this summand contributes to the sum.
If $|a_i^T x x^* - b_i| = m > M$, then choice of w_i such that the quantity:

$$\begin{aligned} g(w_i) &= \frac{m^2}{w_i + 1} + M^2 w_i \\ \frac{d}{dw_i} g(w_i) &= \frac{-m^2}{(w_i + 1)^2} + M^2 = 0 \\ \frac{m}{M} &= w_i + 1 \\ w_i &= \frac{m}{M} - 1 \end{aligned}$$

Evaluating we have that the quantity is: $M(2m - M)$ which is precisely $\phi(a_i^T x - b_i || a_i^T x - b_i| > M)$.

If $|a_i^T x - b_i| < M$, then the w_i which minimizes the quantity is precisely $w_i = 0$, since the derivative is 0 at negative w_i so we check the boundary condition to be the minima for the function. When w_i the quantity is m^2. This is precisely $\phi(a_i^T x - b_i || a_i^T x - b_i| < M)$.

Suppose $(x^\star, w^\star)$ is not the minimum for (a). Then let x' be the optimal value. We con construct w' as follows:

If $|a_i^T(x') - b_i| < M$ then consider $w_i' = 0$.

If $|a_i^T x' - b_i| \geq M$ then $w_i' = \frac{m}{M} - 1$.

Clearly x', w' is a solution more optimal than $(x^\star, w^\star)$ and we have a contradiction. Thus $x^\star$ is optimal for (a).

Now suppose we have an optimal value $x^\star$ for (a). We can construct $w^\star$, and $(x^\star, w^\star)$ is optimal for b. Since otherwise there is some (x', w') with optimal value and since we shown that (x', w') simply takes on the same value as the (a) with x', we have a contradiction.

(b) $b \Leftrightarrow c$:

Let us consider an optimal solution $(x^\star, w^\star)$ from (b). Then we have:

$$\begin{aligned} -u_i - v_i \leq a_i^T x^\star - b_i &\leq u_i + v_i \\ 0 \leq u_i &\leq M \\ v_i &\geq 0 \end{aligned}$$

If $|a_i^T x^\star - b_i| < M$ then we claim that the optimal u_i, v_i is $u_i = m, v_i = 0$.

If $u_i > m$ the smallest v_i possible is $v_i = 0$ and this solution is strictly worse.

If $u_i < m$ the smallest v_i possible is $m - u_i$. We have:

$$\begin{aligned} u_i^2 + 2M(m - u_i) &> m^2 \\ 2M &> m + u_i \end{aligned}$$

So the best solution is when $u_i = m, v_i = 0$. This yields m^2, which is the precise contribution of the ith summand in the sum from (b).

Similarly if $|a_i^T x^\star - b_i| > M$ then $u_i = M$, $v_i = m - M$ is optimal. We do a similar casing strategy as above to argue this claim. The best solution ends up being $M^2 + 2M(m - M)$.

Problem: 4.6 A convex optimization problem can have only *linear* equality constraint functions. In some special cases, however, it is possible to handle convex equality constraint functions, i.e., constraints of the form $h(x) = 0$, where h is convex. We explore this idea in this problem.
Consider the optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h(x) = 0, \end{array} \tag{4.65}$$

where f_i and h are convex functions with domain $\mathbb{R}^n$. Unless h is affine, this is *not* a convex optimization problem.
Consider the related problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h(x) \leq 0, \end{array} \tag{4.66}$$

where the convex equality constraint has been relaxed to a convex inequality. This problem is, of course, convex.
Now suppose we can guarantee that at any optimal solution $x^\star$ of the convex problem (4.66), we have $h(x^\star) = 0$, i.e., the inequality $h(x) \leq 0$ is always active at the solution. Then we can solve the (nonconvex) problem (4.65) by solving the convex problem (4.66).
Show that this is the case if there is an index r such that

- f_0 is monotonically increasing in x_r
- $f_1, \ldots, f_m$ are non decreasing in x_r
- h is monotonically decreasing in x_r

We will see specific examples in exercises **4.31** and **4.58**.

Notes: Can visualize the problem as a ray.

Solution:

Note that the notion of monotonic is strict in BV. Non decreasing and non increasing is non strict. Let us consider an optimal point $x^\star = (x_1^\star, x_2^\star, ..., x_n^\star)$. Then all points on the ray $(x, x_2^\star, ..., x_n^\star)$ satisfy the $f_i(x) \leq 0$ conditions. Additionally we claim that $h(x^\star) = 0$ otherwise, we can improve our objective function an some small enough ball $B_\epsilon(x^\star) \cap R$ (intersected with the ray).

Problem: 4.7 Consider a problem of the form

$$\begin{array}{ll} \text{minimize} & \dfrac{f_0(x)}{c^Tx+d} \\ \text{subject to} & f_i(x) \le 0, \quad i=1,\dots,m \\ & Ax=b \end{array}$$

where $f_0, f_1, \dots, f_m$ are convex, and the domain of the objective function is defined as $\{x \in \text{dom } f_0 \mid c^Tx + d > 0\}$.

1. Show that this is a quasiconvex optimization problem.
2. Show that the problem is equivalent to

$$\begin{array}{ll} \text{minimize} & g_0(y,t) \\ \text{subject to} & g_i(y,t) \le 0, \quad i=1,\dots,m \\ & Ay=bt \\ & c^Ty+dt=1, \end{array}$$

where g_i is the perspective of f_i (see §3.2.6). The variables are $y \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Show that this problem is convex.

3. Following a similar argument, derive a convex formulation for the *convex-concave* fractional problem

$$\begin{array}{ll} \text{minimize} & \dfrac{f_0(x)}{h(x)} \\ \text{subject to} & f_i(x) \le 0, \quad i=1,\dots,m \\ & Ax=b \end{array}$$

where $f_0, f_1, \dots, f_m$ are convex, h is concave, the domain of the objective function is defined as $\{x \in \text{dom } f_0 \cap \text{dom } h \mid h(x) > 0\}$ and $f_0(x) \ge 0$ everywhere.

As an example, apply your technique to the (unconstrained) problem with

$$f_0(x) = \frac{\text{tr}(F(x))}{m}, \quad h(x) = (\det(F(x)))^{1/m},$$

with $\text{dom}(f_0/h) = \{x \mid F(x) \succ 0\}$, where $F(x) = F_0 + x_1F_1 + \cdots + x_nF_n$ for given $F_i \in \mathbb{S}^m$. In this problem, we minimize the ratio of the arithmetic mean over the geometric mean of the eigenvalues of an affine matrix function $F(x)$.

Solution:

(a) We need to show that $\frac{f_0(x)}{c^Tx+d}$ is quasi convex. From Chapter 3, subsection, "Properties of quasi convex functions", it suffices to show that the function satisfies Jensen's inequality for quasi convex functions. In particular we want to show for any $x, y \in \textbf{dom} f$:

$$\begin{aligned} f(\theta x + (1-\theta)y) &\le \max(f(x), f(y)) \\ \frac{f_0(\theta x + (1-\theta)y)}{c^T(\theta x + (1-\theta)y) + d} &\le \frac{\theta f_0(x) + (1-\theta)f_0(y)}{c^T(\theta x + (1-\theta)y) + d} && \text{By convexity} \\ \text{WLOG, f(x) ¡ f(y). suffices to show:} & \\ \frac{\theta f_0(x) + (1-\theta)f_0(y)}{c^T(\theta x + (1-\theta)y) + d} &\le \frac{f_0(x)}{c^Tx+d} \\ (\theta c^Tx + \theta d)f_0(x) + ((1-\theta)c^Tx + (1-\theta)d)f_0(y) &\le (c^T\theta x + c^T(1-\theta)y + \theta d + (1-\theta)d)f_0(x) \\ (1-\theta)(c^Tx+d)f_0(y) &\le (1-\theta)(c^Ty+d)f_0(x) \\ \frac{f_0(y)}{c^Ty+d} &\le \frac{f_0(x)}{c^Tx+d} \\ f(y) &\le f(x) \end{aligned}$$

(b) TODO.

Problem: 4.8 Give an explicit solution of each of the following LPs.

(a) *Minimizing a linear function over an affine set.*

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b. \end{array}$$

(b) *Minimizing a linear function over a halfspace.*

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a^T x \leq b, \end{array}$$

where $a \neq 0$.

(c) *Minimizing a linear function over a rectangle.*

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & l \preceq x \preceq u, \end{array}$$

where l and u satisfy $l \preceq u$.

(d) *Minimizing a linear function over the probability simplex.*

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{1}^T x = 1, \quad x \succeq 0. \end{array}$$

What happens if the equality constraint is replaced by an inequality $\mathbf{1}^T x \leq 1$?

We can interpret this LP as a simple portfolio optimization problem. The vector x represents the allocation of our total budget over different assets, with x_i the fraction invested in asset i. The return of each investment is fixed and given by $-c_i$, so our total return (which we want to maximize) is $-c^T x$. If we replace the budget constraint $\mathbf{1}^T x = 1$ with an inequality $\mathbf{1}^T x \leq 1$, we have the option of not investing a portion of the total budget.

(e) *Minimizing a linear function over a unit box with a total budget constraint.*

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{1}^T x = \alpha, \quad 0 \preceq x \preceq 1, \end{array}$$

where α is an integer between 0 and n. What happens if α is not an integer (but satisfies $0 \leq \alpha \leq n$)? What if we change the equality to an inequality $\mathbf{1}^T x \leq \alpha$?

(f) *Minimizing a linear function over a unit box with a weighted budget constraint.*

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & d^T x = \alpha, \quad 0 \preceq x \preceq 1, \end{array}$$

with $d \succ 0$, and $0 \leq \alpha \leq \mathbf{1}^T d$.

Notes:

- Good linear algebra problems all exercises are good to redo.

Solution:

(a) First if $Ax = b$ is not feasible then the optimal value is ∞.

Now suppose $Ax = b$ is feasible. There are two cases:

If c is orthogonal to the null space. Then c has two components:

$$c = c_{\mathcal{N}_\perp} + c_{\mathcal{N}}$$

We have $Ac_{\mathcal{N}} = 0$. Additionally by fundamental theorem of linear algebra, $c_{\mathcal{N}_\perp}$ must lie in the range of A^T. So we have $c = A^T c' + c_{\mathcal{N}}$.

Suppose that $c_{\mathcal{N}} = 0$. Then we have:

$$\begin{aligned} c^T x &= c'^T Ax \\ &= c'^T b \end{aligned}$$

For c' that is the orthogonal component of c to the null space of A.

If c is not orthogonal to the null space then we have that $c_{\mathcal{N}} \neq 0$. Let us consider, $x + -tc_{\mathcal{N}}$:

$$c'^T Ax + c_{\mathcal{N}}^T x - c'^T Atc_{\mathcal{N}} - tc_{\mathcal{N}}^T c_{\mathcal{N}}$$

Which is unbounded below as $t \to \infty$.

(b) Similarly to above it is best to case when decomposing into orthogonal and parallel pieces with respect to a subspace. We can write $c = a\lambda + a_\perp$ and case.

If $\lambda_\perp \neq 0$ then intuitively we can push the optimum to $-\infty$. Consider $x = -ta_\perp - aM$, with $M = \frac{2b}{\|a\|_2^2}$:

$$\begin{aligned} a^T(ta_\perp - aM) &= -a^T aM < b \\ M &> \frac{b}{\|a\|_2^2} \end{aligned}$$

And we have:

$$(a\lambda + a_\perp\lambda_\perp)^T(-ta_\perp - aM) = -\|a\|^2 M - t\|a_\perp\|$$

If $\lambda_\perp = 0$, then we case on λ. If $\lambda > 0$:

$c = a\lambda$, let $x = -ta$ where $t \to \infty$, we have that:

$$-ta^T a < b \qquad \text{for large enough t}$$

And for the objective function:

$$(a\lambda)^T(-ta) = -t\lambda\|a\|_2^2 \to -\infty$$

For $\lambda = 0$, the problem is trivial.

For $\lambda < 0$:

It is helpful to consider $x = ta + t_2 v$ where v is perpendicular to a. We will observe that t_2 can vary without affecting the below analysis. We thus will consider $t_2 = 0$. We have to optimize:

$$t\lambda\|a\|_2^2$$

With the constraint:

$$\begin{aligned} t\|a\|_2^2 &\leq b \\ t &\leq \frac{b}{\|a\|_2^2} \end{aligned}$$

To minimize the objective we choose the largest $t = \frac{b}{\|a\|_2^2}$. Thus the objective has value: λb.

(c) **My approach:**

Note we can simply consider the vertices of the rectangle as candidates for an optimum. If we visualize $c^T x \leq t$ as a hyperplane, as we lower t, we intersect eventually only with a vertex or a face of the box. Clearly the optimum does not lie in the interior of the box, since if it did we can give x a directional shift in the c^T direction to improve our minimum. So the minimum lies on the boundary, and it suffices to check vertices only.

Alt approach:

Consider the problem component wise:

minimize:

$\sum c_i x_i$

subject to:

$l_i \leq x \leq u_i$

If $c_i > 0$ then clearly l_i optimizes this summand so $x_i = l_i$.

If $c_i < 0$ then clearly u_i optimizes this summand so $x_i = u_i$.

If $c_i = 0$, then any x_i suffices.

(d) This part, no need to over complicate things. Let us consider $c_1, c_2, ..., c_n$ and WLOG, suppose that $c_1 \leq c_2 \leq ... \leq c_n$. Clearly the optimal value is lower bounded by c_1. We can achieve this by considering a probability vector e_1. In the inequality case we consider $\min c_1, 0$ to be our optimal value since we can choose not to invest, if the cost is negative.

(e) WLOG let consider the costs $c_1, c_2, ..., c_n$, WLOG suppose that $c_1 \leq c_2 \leq ... \leq c_n$. Then we choose the α smallest costs.

If α is not an integer, we choose the $\lfloor \alpha \rfloor$ smallest costs and $\{\alpha\}$ of the $\lfloor \alpha \rfloor + 1$ cost.

(f) Let us consider new variables $y_i = d_i x_i$. Then we want to minimize:

$\sum c_i (y_i / d_i)$

Subject to:

$0 \preceq y \preceq d$

$d^T x = \alpha, \sum d_i x_i = \alpha$

Let us consider an ordering of: $\frac{c_i}{d_i}$. For the smallest values we assign the maximum value of x_i. For the $\lfloor \alpha \rfloor$ smallest values we assign $x_i = 1$ and for the $\lfloor \alpha \rfloor + 1$ smallest value, we assign $x_i = \{\alpha\}$.

Problem: 4.9 *Square LP.* Consider the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

with A square and non-singular. Show that the optimal value is given by

$$p^\star = \begin{cases} c^T A^{-1} b & A^{-T} c \preceq 0 \\ -\infty & \text{otherwise.} \end{cases}$$

Notes: Good tricky problem.

Solution:

Minimize $c^T x$, we can rewrite this as:

$$(A^{-T} c)^T A x$$

If $(A^{-T} c)^T \preceq 0$ then we simply want to maximize Ax in the $\|\|_1$ sense. So we can consider the bound $Ax \preceq b$, and just use $x = A^{-1} b$. Therefore the optimal value: $c^T A^{-1} b$.

Otherwise one of the values $(A^{-T}c)^T > 0$. Since Ax is full rank we can lower this entry as low as we want.

Problem: 4.10 Work out the details on page 147 of 84.3.
Explain in detail the relation between:

- The feasible sets
- The optimal solutions
- The optimal values

of the standard form LP and the original LP.

Notes:

Nice application of slack variables. What are the general LP –¿ slack LP –¿ standard LP forms?

Solution:

We essentially want to show the equivalence between three LPs. The original LP, the LP with slack variables, and the standard LP form.

First let us show the equivalence between the original LP and the LP with slack variables.

Recall the original LP as:

minimize:

$c^T x + d$

subject to:

$Gx + s \leq h, Ax = b$

The LP with slack variables (recall that slack variables allow us to go from inequality to equals) we have:

minimize:

$c^T x + d$

subject to:

$Gx + s = h, Ax = b, s \succeq 0.$

First let us describe the feasible sets. Let x be feasible in the original LP. Then this induces $s = h - Gx$, (x, s) as a feasible point in the slack problem. Note that $s \succeq 0$ since $Gx \preceq h$. Similarly let (x, s) be in the feasible set of the slack problem. Clearly x is also in the feasible set of the original problem.

Let $(x^\star, s^\star)$ be an optimal solution for the slack problem. We claim $x^\star$ is optimal for the original problem. Suppose there is x' which has $f(x') <$ and x' is feasible. Then this induces (x', s'), but $f(x') < f(x^\star)$ for the slack solution, which is a contradiction. We make a similar argument as to why optimal in original implies optimal in slack. Therefore the optimal value of both problems re the same. The optimal solutions induce each other as well, there is a bijection $x^\star \mapsto (x^\star, s^\star)$, with $s^\star = h - Gx^\star$.

Standard to Slack:

First let us analyze the feasible sets. Let us consider an element (x^+, x^-, s). We have that $(x^+ + x^-, s)$ is a solution to the slack problem. Now consider (x, s) a solution to the slack problem. We write $x = x^+ - x^-$ where if $x_i \geq 0$ then $x_i^+ = x_i, x_i^- = 0$ and if $x_i < 0$ then $x_i^+ = 0, x_i^- = x_i$.

For optimal values argument is similar to the first part.

Problem: 4.11 Formulate the following problems as LPs. Explain in detail the relation between the optimal solution of each problem and the solution of its equivalent LP.

(a) Minimize $\|Ax - b\|_\infty$ (ℓ_∞-norm approximation).

(b) Minimize $\|Ax - b\|_1$ (ℓ_1-norm approximation).

(c) Minimize $\|Ax - b\|_1$ subject to $\|x\|_\infty \leq 1$.

(d) Minimize $\|x\|_1$ subject to $\|Ax - b\|_\infty \leq 1$.

(e) Minimize $\|Ax - b\|_1 + \|x\|_\infty$.

In each problem, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ are given. (See §6.1 for more problems involving approximation and constrained approximation.)

Notes: First two are good to redo. After that it is very similar.

(a) Note that for $Ax - b$ we are choosing the component with the smallest magnitude. So it makes sense to check component-wise for the constraints and minimize over some length:

minimize: t

subject to: $-t \leq a_i x - b_i \leq t$, $i \in [n]$

Let $x^\star$ be an optimal value for the original problem $\|Ax^\star - b\|_\infty$. Let $t = \|Ax^\star - b\|_\infty$. Clearly $x^\star$ satisfies all bounds. If there exists $t' < \|Ax^\star - b\|_\infty$, then consider x' and we have a contradiction, since $\|Ax' - b\|_\infty \leq t < \|Ax^\star - b\|_\infty$.

(b) Note for $Ax - b$ for each component we are measuring the distance from 0. There are essentially two cases, if $ax_i - b_i \geq 0$ and $ax_i - b_i < 0$. So we have to consider these separately in the LP:

minimize: $1^T t + 1^T s$

subject to: $ax_i - b_i + t_i \geq 0, t_i \geq 0$, $ax_i - b_i - s_i \leq 0, s_i \geq 0$.

If $ax_i - b_i \geq 0$ then $t_i = 0$, but the optimal $s_i = |ax_i - b_i|$. If $ax_i - b_i < 0$, then $t_i = |ax_i - b_i|$ and $s_i = 0$.

Let $x^\star$ be optimal. Then $\|Ax^\star - b\|_1 = \sum_{i=1}^n |a_i x^\star - b_i|$. Then note for each $a_i x^\star - b_i$ either s_i or t_i takes on this value and the other value is 0. So the contribution to our objective is: $|a_i x^\star - b_i|$.

Let us consider an optimal $(x^\star, t^\star, s^\star)$ which achieves a better objective value than $\|Ax^\star - b\|_1$, which will lead to a contradiction to the original problem.

Alternatively we can consider a cleaner solution:

minimize: $1^T s$

constraints: $s \preceq Ax - b \preceq s$

- Same as b but with the constraint $-1 \preceq x \preceq 1$.

- Consider:

  minimize: $1^T s$

  constraints: $-s \preceq x \preceq s, s \succeq 0$, $-1 \preceq Ax - b \preceq 1$.

  Suppose $x^\star$ is a minimizer. Then the optimal $s^\star = \|x^\star\|_1$. Component wise $s_i = |x_i|$. Since $x^\star$ satisfies $\|Ax - b\|_\infty \leq 1$ it satisfies this constraint in this new problem too. If there is a more optimal $s^\star$ then we will have a contradiction, since the induced x' will have a smaller $\|x'\|_1 \leq 1^T s$.

- minimize: $1^T s + t$

  constraints: $-s \preceq Ax - b \preceq s, -t1 \preceq x \preceq t1$.

  Very similar arguments to above. Not too complicated.

Problem: 4.12
Consider a network of n nodes, with directed links connecting each pair of nodes. The variables in the problem are the flows on each link: x_{ij} will denote the flow from node i to node j. The cost of the flow along the link from node i to node j is given by $c_{ij}x_{ij}$, where c_{ij} are given constants. The total cost across the network is

$$C = \sum_{i,j=1}^{n} c_{ij}x_{ij}.$$

Each link flow x_{ij} is also subject to a given lower bound l_{ij} (usually assumed to be nonnegative) and an upper bound u_{ij}.
The external supply at node i is given by b_i, where:

- $b_i > 0$ means an external flow enters the network at node i
- $b_i < 0$ means that at node i, an amount $|b_i|$ flows out of the network

We assume that $\mathbf{1}^T b = 0$, *i.e.*, the total external supply equals total external demand. At each node we have conservation of flow: the total flow into node i along links and the external supply, minus the total flow out along the links, equals zero.
The problem is to minimize the total cost of flow through the network, subject to the constraints described above.

Notes:

Easy

Solution:

minimize:

$$\sum_{i,j=1}^{n} c_{i,j}x_{i,j}$$

subject to:

$$\ell_{ij} \leq x_{ij} \leq u_{ij}$$
$$\sum_{i\neq j} x_{ij} + b_i + \sum_{i\neq j} x_{ji} = 0, i \in [n]$$

Problem: 4.13 Consider the problem, with variable $x \in \mathbb{R}^n$,

$$\text{minimize} \quad c^T x$$

subject to $Ax \preceq b$ for all $A \in \mathcal{A}$,
where $\mathcal{A} \subseteq \mathbb{R}^{m\times n}$ is the set

$$\mathcal{A} = \left\{ A \in \mathbb{R}^{m\times n} | \overline{A}_{ij} - V_{ij} \leq A_{ij} \leq \overline{A}_{ij} + V_{ij},\, i = 1, \ldots, m,\, j = 1, \ldots, n \right\}.$$

The matrices $\bar{A}$ and V are given. This problem can be interpreted as an LP where each coefficient of A is only known to lie in an interval, and we require that x must satisfy the constraints for all possible values of the coefficients. Express this problem as an LP. The LP you construct should be efficient, *i.e.*, it should not have dimensions that grow exponentially with n or m.

Solution:

First note that due to this inequality $\overline{A}_{ij} - V_{ij} \leq A_{ij} \leq \overline{A_{ij}} + V_{ij}$, V_{ij} must be non negative. Analyzing the constraint we have that: $Ax \preceq b$ for all $A \in \mathcal{A}$. The maximum value $a_1 x$ can be, for a fixed x, can be analyzed component wise,

$\overline{a}_1 x + \sum_{i=1}^{n} v_{1i}|x_i| = \overline{a}_1 x + v_i|x|$. This quantity is achieved for some parameters of a_i. Therefore this quantity must be less than b_i. We can combine all equations into the form:

$$\overline{A}x + V|x| \preceq b$$

Now we can consider making this into standard LP form, by introducing a second variable to bound the absolute value.

$$\overline{A}x + Vy \preceq b, -y \preceq x \preceq y$$

Note that y has feasible values since y_i can take on the value $|x_i|$. Therefore any x feasible in the original problem is feasible in our equation. Suppose (x, y) is feasible in the second formulation. Then clearly $(x, |x|)$ is feasible as well. Therefore we satisfy the original formulation. Therefore any optimal $x^\star$ will be optimal in the second solution and vice versa.

Problem: 4.14 *Approximating a matrix in infinity norm.* The ℓ_∞-norm induced norm of a matrix $A \in \mathbf{R}^{m\times n}$, denoted $\|A\|_\infty$, is given by

$$\|A\|_\infty = \sup_{x\neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}|.$$

This norm is sometimes called the max-row-sum norm, for obvious reasons (see §A.1.5). Consider the problem of approximating a matrix, in the max-row-sum norm, by a linear combination of other matrices. That is, we are given $k+1$ matrices $A_0, \dots, A_k \in \mathbf{R}^{m\times n}$, and need to find $x \in \mathbf{R}^k$ that minimizes

$$\|A_0 + x_1 A_1 + \cdots + x_k A_k\|_\infty .$$

Express this problem as a linear program. Explain the significance of any extra variables in your LP. Carefully explain how your LP formulation solves this problem, *e.g.*, what is the relation between the feasible set for your LP and this problem?

Solution:

Note we can do a "brute force" style bounding on each of the individual terms in the m sums that are considered for the max. This is because an absolute value term can be conveniently bounded by one variable as a lower and upper bound. So considering the matrix infinity norm:

$$\|A_0 + x_1 A_1 + ... + x_k A_k\|_\infty = \max_{i=1,\dots,m} \sum_{j=1}^{n} |a_{ij}^0 + x_1 a_{ij}^1 + ... + x_k a_{ij}^k|$$

First we can consider the min of y, which will be a variable in which we attempt to represent as the infinity norm.

Each term in a summand can be represented as:

$$-t_{ij} \le a_{ij}^0 + x_1 a_{ij}^1 + ... + x_k a_{ij}^k \le t_{ij}$$

For a total of $2nm$. Then we have $y \ge \sum_{j=1}^{n} t_{ij}, i = 1, .., n$. We claim that this linear program will solve for $x_1, .., x_k$ which minimizes the matrix infinity norm. Suppose $x^\star$ is optimal for the original problem. So $f(x^\star) = y^\star$. Clearly if $y^{\star'}$ to our optimal problem is $< y^\star$ we have a contradiction since we can simply consider an induced $x^{\star'}$. Otherwise if $y^{\star'} > y^\star$ we can consider optimal $x^\star$ in the linear program to improve upon this value.

Note that the solution manual "visually" reduces the system of $2nm$ into a generalized inequality:

$$-S \preceq_K A_0 + A_1 x_1 + ... + A_k x_k \preceq_K S, y \geq \sum_{j=1}^{n} s_{ij}, i = 1, ..., m$$

Where K is the cone, $\{X | X_{ij} \geq 0\}$.

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

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x_i \in \{0,1\}, \quad i = 1, \ldots, n. \end{array} \tag{4.67}$$

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 $0 \leq x_i \leq 1$:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & 0 \leq x_i \leq 1, \quad i = 1, \ldots, n. \end{array} \tag{4.68}$$

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?

Notes: Very very chill problem.

Solution:

1. The optimal value in the relaxed LP is a lower bound on the optimal in the boolean LP. Consider optimal $x_b^\star$. Clearly $x_b^\star$ satisfies the constraints of the relaxed LP. So the optimal value of the relaxed LP is upper bounded by $c^T x_b^\star$.

2. If the LP relaxation has a solution $x_i \in \{0,1\}$ then both problems have the same optimal value.

Problem: 4.16 We consider a linear dynamical system with state $x(t) \in \mathbb{R}^n$, $t = 0, \ldots, N$, and actuator or input signal $u(t) \in \mathbb{R}$, for $t = 0, \ldots, N-1$. The dynamics of the system is linear recurrence

$$x(t+1) = Ax(t) + bu(t), \quad t = 0, \ldots, N-1, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$ are given. We assume that the initial state is zero, i.e., $x(0) = 0$.
The *minimum fuel optimal control problem* is to choose the inputs $u(0), \ldots, u(N-1)$ so as to minimize the total fuel consumed, which is given by

$$F = \sum_{t=0}^{N-1} f(u(t)), \tag{2}$$

subject to the constraint that $x(N) = x_{\text{des}}$, where N is the (given) time horizon, and $x_{\text{des}} \in \mathbb{R}^n$ is the (given) desired final or target state. The function $f : \mathbb{R} \to \mathbb{R}$ is the *fuel use map* for the actuator, and gives the amount of fuel used as a function of the actuator signal amplitude. In this problem we use

$$f(a) = \begin{cases} |a| & |a| \leq 1 \\ 2|a| - 1 & |a| > 1. \end{cases} \tag{3}$$

This means that fuel use is proportional to the absolute value of the actuator signal, for actuator signals between -1 and 1; for larger actuator signals the marginal fuel efficiency is half.
Formulate the minimum fuel optimal control problem as an LP.

Notes:

Developed alt solution. Book solution very clean. Should review it.

Solution: First we can formulate $x(t+1) = Ax(t) + bu(t), t = 0, ..., N-1$ as a system of equations with respect to $u(t)$.

For $x(0) = 0, x(1) = bu(0)$.

$$\begin{aligned} x(2) &= Abu(0) + bu(0) \\ x(3) &= A^2bu(0) + Abu(0) + bu(1) \\ &\vdots \\ x(N) &= A^{N-1}bu(0) + A^{N-2}bu(0) + \sum_{i=1}^{N-2} A^{i-1}bu(N-1-i) \end{aligned}$$

Now we have a linear equality for the distant.

$$A^{N-1}bu(0) + A^{N-2}bu(0) + \sum_{i=1}^{N-2} A^{i-1}u(N-1-i) = x_{\text{des}}$$

Now to model the sum of the $N-1$ functions we just have to model each function individually. Let $x_t = u(t)$.

Since we dealing with casework and absolute values, I found it to be most natural to use two sets of two variables (two sets for the cases) (two variables for the + and - cases of the absolute value). Namely let us consider the variables:

$$\begin{aligned} x_t + z_t^+ &\geq 0 & x_t - z_t^- &\leq 0 & z_t^+ &\geq 0 & z_t^- &\geq 0 \\ -x_t + w_t^+ &\geq 1 & -x_t - w_t^- &\leq -1 & w_t^+ &\geq 0 & w_t^- &\geq 0 \end{aligned}$$

This system of equations seems somewhat unmotivated: the right inequalities are standard we want to make our latent variables ≥ 0. The top left two if we consider the sum $z_t^+ + z_t^-$, this will always be equal to $|x_t|$ in the minimization case. Now the bottom left two I had to case and experiment to develop them. First if $|x_t| \leq 1$ then $w_t^+ + w_t^-$ will always equal 2. If $0 \leq x_t \leq 1$ then $w_t^+ = 1 + x_t$ and $w_t^- = 1 - x_t$. If $-1 \leq x_t \leq 0$ then w_t

If $|x_t| \geq 1$ then $w_t^+ = x_t + 1$ and $w_t^- = 0$. Similarly if $|x_t| \leq 1$ then $w_t^- = x_t + 1, w_t^+ = 0$. We wish to minimize: $z_t^+ + z_t^- + w_t^+ + w_t^- - 2$. If $|x_t| \leq 1$ then the minimal value will be $|x_t|$. If $|x_t| \geq 1$ then the minimal value will be $|x_t| + |x_t| - 1$ which in either case matches the function.

So we have a LP! Showing equivalence is not hard.

The book as a much cleaner solution. The distance equality is the same. However since the function $f(u)$ is essentially the max of two functions $|u|, 2|u| - 1$, we can consider the following: minimize $1^T t$, $-y \preceq u \preceq y$ (typical absolute bounding), $t \succeq y, t \succeq 2y - 1$.

Problem: 4.17 Optimal activity levels. We consider the selection of n nonnegative activity levels, denoted $x_1, \ldots, x_n$. These activities consume m resources, which are limited. Activity j consumes A_{ij} of resource i, where $A_{ij} \geq 0$ are given. The total resource consumption is additive; the total of resource i consumed is $c_i = \sum_{j=1}^n A_{ij} x_j$. (Ordinarily we have $A_{ij} \geq 0$, i.e., activity j consumes resource i as a product. But allow the possibility that $A_{ij} < 0$, which means that activity j actually generates resource i as a by-product.) Each resource consumption is limited: we must have $c_i \leq c_i^{\max}$, where $c_i^{\max}$ are given. Each activity generates revenue, which is a piecewise-linear concave function of the activity level:

$$r_j(x_j) = \begin{cases} p_j x_j & 0 \leq x_j \leq q_j \\ p_j q_j + p_j^{\text{disc}}(x_j - q_j) & x_j \geq q_j. \end{cases} \tag{4}$$

Here $p_j > 0$ is the basic price, $q_j > 0$ is the quantity discount level, and p_j^{disc} is the quantity discount price, for (the product of) activity j. (We have $0 < p_j^{\text{disc}} < p_j$.) The total revenue is the sum of the revenues associated with each activity, i.e., $\sum_{j=1}^n r_j(x_j)$. The goal is to choose activity levels that maximize total revenue, subject to the given resource limits. Show how to formulate this problem as an LP.

Notes: Slightly different flavor than above ones since this is a LP Max problem.

Solution:

First we have the constraint on the variable x:

$Ax \preceq c^{\max}$.

Now we want to maximize a quantity since we are attempting to maximize revenue. So we can begin by considering some latent variable s, and finding constraints which allow $1^T s$ to represent the revenue.

Usually for minimization problems we had s as a lower $-s$ and upper bound s with respect to x. But in this case, this does not make sense since if we are maximizing s, we can just make s very large in all dimensions. So we can allow each dimension of s to reach the max of the two possible revenue outputs. Note that the revenue function is essentially the max of $p_j x_j$ and $p_j q_j + p_j^{\text{disc}}(x_j - q_j)$. Therefore we can consider:

$$s \preceq \begin{pmatrix} p_1 x_1 \\ p_2 x_2 \\ \vdots \\ p_n x_n \end{pmatrix} \qquad s \preceq \begin{pmatrix} p_1 q_1 + p_1^{\text{disc}}(x_1 - q_1) \\ \vdots \\ p_n q_n + p_n^{\text{disc}}(x_n - q_n) \end{pmatrix}$$

Showing equivalence is not hard. Will skip for now.

Problem: 4.18 Separating hyperplanes and spheres. Suppose you are given two sets of points in $\mathbb{R}^n$, $\{v^1, v^2, \dots, v^K\}$ and $\{w^1, w^2, \dots, w^L\}$. Formulate the following two problems as LP feasibility problems.

(a) Determine a hyperplane that separates the two sets, i.e., find $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$ with $a \neq 0$ such that

$$a^T v^i \leq b, \quad i = 1, \dots, K, \tag{5}$$
$$a^T w^i \geq b, \quad i = 1, \dots, L. \tag{6}$$

Note that we require $a \neq 0$, so you have to make sure that your formulation excludes the trivial solution $a = 0$, $b = 0$. You can assume that

$$\operatorname{rank} \begin{bmatrix} v^1 & v^2 & \cdots & v^K & w^1 & w^2 & \cdots & w^L \\ 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \end{bmatrix} = n + 1 \tag{7}$$

(i.e., the affine hull of the $K + L$ points has dimension n).

(b) Determine a sphere separating the two sets of points, i.e., find $x_c \in \mathbb{R}^n$ and $R \geq 0$ such that

$$\|v^i - x_c\|_2 \leq R, \quad i = 1, \dots, K, \tag{8}$$
$$\|w^i - x_c\|_2 \geq R, \quad i = 1, \dots, L. \tag{9}$$

(Here x_c is the center of the sphere; R is its radius.)

(See chapter 8 for more on separating hyperplanes, separating spheres, and related topics.)

Note:

Part B is a nice problem. How to approach LP with quadratics?

Solution:

(a) First clearly the hyperplane half-planes are themselves linear constraints. We simply consider $K+L$ sets of equations:

$$a^T v_i \leq b, i = 1, .., K \qquad\qquad a^T w_i \geq b, i = 1, .., L$$

Referencing the book, we can consider this as one matrix inequality:

$$Bx \succeq 0$$

Where $B = \begin{pmatrix} {v^1}^T & -1 \\ \vdots & \vdots \\ {w^L}^T & -1 \end{pmatrix}$ Since B has rank $n+1, K+L \geq n+1$.

We can extract any non trivial x by adding another small constraint. Since technically $x = 0$ is a solution to the problem, but it is an uninteresting one. So we can consider $1^T Bx = 1$. Note that there should exists such an x since Bx has full rank, we can just consider any x (non trivial) which satisfies $Bx \succeq 0$, and scale it down to satisfy our new constraint $1^T Bx = 1$.

(b) Had to reference the book for this one. But the approach here is to isolate the linear variables and replace all of the quadratics with a latent variable and treat it as a linear system. We can backtrack to solve for the quadratic terms once we find a feasible solution and value for the latent variable. The latent variable is quadratic and depends on R, x_c and as long as one of the variables we solve for is R or x_c in the LP then we can backtrack to solve for R, x_c (note since we will also have solved for the latent).

$$\|w^i\|_2^2 + \|x_c\|_2^2 - 2(w^i)^T x_c \le R^2$$
$$\|v^i\|_2^2 + \|x_c\|_2^2 - 2(v^i)^T x_c \ge R^2$$
$$\|w^i\|_2^2 - 2(w^i)^T x_c \le R^2 - \|x_c\|_2^2 = q$$
$$\|v^i\|_2^2 - 2(v^i)^T x_c \ge R^2 - \|x_c\|_2^2 = q$$

Minimize q, subject to the above constraints (all linear). Once we solve for some q, x_c then we can backtrack to find R. Arguing feasibility is easy. Clearly if there exists such a R and x_c then these values work for q, and our LP will at least output this solution (it will try to find a "smaller" q in the $|q|$ sense). If such a q, x_c exist then we can backtrack to find R, and we are done.

Problem: 4.19 Consider the problem

$$\begin{aligned} \text{minimize} \quad & \|Ax - b\|_1/(c^T x + d) && (10) \\ \text{subject to} \quad & \|x\|_\infty \le 1, && (11) \end{aligned}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, and $d \in \mathbb{R}$. We assume that $d > \|c\|_1$, which implies that $c^T x + d > 0$ for all feasible x.

(a) Show that this is a quasiconvex optimization problem.

(b) Show that it is equivalent to the convex optimization problem

$$\begin{aligned} \text{minimize} \quad & \|Ay - bt\|_1 && (12) \\ \text{subject to} \quad & \|y\|_\infty \le t && (13) \\ & c^T y + dt = 1, && (14) \end{aligned}$$

with variables $y \in \mathbb{R}^n$, $t \in \mathbb{R}$.

Solution:

(a) Here we use Jensen's inequality for quasiconvex functions. We have to show that $x, y \in \textbf{dom f}, \theta \in [0,1]$ we have:

$$f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$
$$\frac{\|A(\theta x + (1-\theta)y) - b\|_1}{c^T(\theta x + (1-\theta)y) + d} \le \frac{\|Ax - b\|_1}{c^T x + d} \qquad \text{assume } f(x) \ge f(y)$$
$$\text{by triangle inequality}$$
$$(\theta\|Ax - b\| + (1-\theta)\|Ay - b\|)(c^T x + d) \le \|Ax - b\|[\theta(c^T x + d) + (1-\theta)(c^T y + d)]$$
$$(1-\theta)\|Ay - b\|(c^T x + d) \le \|Ax - b\|(1-\theta)(c^T y + d)$$

Which is true by assumption of $f(x) \ge f(y)$.

(b) Let us first rewrite the new convex optimization problem. Assuming $t \ne 0$ we can get motivation to manipulate the problem:

minimize: $t\|A\frac{y}{t} - b\|_1$

subject to: $\|\frac{y}{t}\|_\infty \le 1$ and with $c^T(\frac{y}{t}) + d = \frac{1}{t}$.

Rewriting the min as $\frac{\|A\frac{y}{t} - b\|_1}{1/t}$ and by the constraint: $\frac{\|A\frac{y}{t} - b\|_1}{c^T(\frac{y}{t}) + d}$.

Now let us consider some optimal $x^\star$ in the original equation. Then we have that $c^T x^\star + d = q^\star$. We have that $t = \frac{1}{q^\star}$. We also have that $y = \frac{x^\star}{q^\star}$. Note that $q^\star > 0$ by assumption. This is clearly feasible. Suppose there is an optimal $y^\star, t^\star$ which yields a smaller value. Then we consider $x = \frac{y^\star}{t^\star}$ for a contradiction. Note that all of this insight was possible since we were able to manipulate the new LP.

Problem: 4.20 *Power assignment in a wireless communication system.* We consider n transmitters with powers $p_1, \dots, p_n \geq 0$, transmitting to n receivers. These powers are the optimization variables in the problem. We let $G \in \mathbb{R}^{n\times n}$ denote the matrix of *path gains* from the transmitters to the receivers; $G_{ij} \geq 0$ is the path gain from transmitter j to receiver i. The *signal power* at receiver i is then $S_i = G_{ii}p_i$, and the *interference power* at receiver i is $I_i = \sum_{k\neq i} G_{ik}p_k$. The *signal to interference plus noise ratio*, denoted SINR, at receiver i, is given by $S_i/(I_i+\sigma_i)$, where $\sigma_i > 0$ is the (self-) noise power in receiver i. The objective in the problem is to maximize the minimum SINR ratio, over all receivers, *i.e.*, to maximize

$$\min_{i=1,\dots,n} \frac{S_i}{I_i + \sigma_i}.$$

There are a number of constraints on the powers that must be satisfied, in addition to the obvious one $p_i \geq 0$. The first is a maximum allowable power for each transmitter, *i.e.*, $p_i \leq P_i^{\max}$, where $P_i^{\max} > 0$ is given. In addition, the transmitters are partitioned into groups, with each group sharing the same power supply, so there is a total power constraint for each group of transmitter powers. More precisely, we have subsets $K_1, \dots, K_m$ of $\{1, \dots, n\}$ with $K_1 \cup \dots \cup K_m = \{1, \dots, n\}$, and $K_j \cap K_l = 0$ if $j \neq l$. For each group K_l, the total associated transmitter power cannot exceed $P_l^{\mathrm{BP}} > 0$:

$$\sum_{k\in K_l} p_k \leq P_l^{\mathrm{BP}}, \quad l = 1, \dots, m.$$

Finally, we have a limit $P_i^{\mathrm{rc}} > 0$ on the total received power at each receiver:

$$\sum_{k=1}^{n} G_{ik}p_k \leq P_i^{\mathrm{rc}}, \quad i = 1, \dots, n.$$

(This constraint reflects the fact that the receivers will saturate if the total received power is too large.)
Formulate the SINR maximization problem as a generalized linear-fractional program.

Notes:

Looks intimidating but easy to break down to LP.

Solution:

Conveniently the minimum SINR ratio over all receivers is almost in generalized linear fractional form. We have that $S_i = G_{ii}p_i$. And that the denominator can be viewed as: $a_ip + \sigma_i$ where a_i is $[G_{i1}, ..., 0, ..., G_{in}]$, where we have a zero at $k = i$. The numerator can be thought of as b_ip where b_i is zeros everywhere except at the ith entry with G_{ii}.

In terms of constraint we have $\sum p_k \leq p = P_l^{gp}$ which are just collection of p_i and bounding by constant. Similarly the total received power is a weighted sum of all p bounded by a constant.