Some Solutions to Homework due 2/16

1.3, #30. Suppose that $S_1, \dots, S_m$ are all convex subsets of $\boldsymbol{R}^n$ (for the same n). To prove: $S_1 \cap \cdots \cap S_m$ is convex.

Proof: Suppose that $\boldsymbol{x}, \boldsymbol{y} \in S_1 \cap \cdots \cap S_m$. We must show that any point $\boldsymbol{v}$ on the line segment joining $\boldsymbol{x}$ and $\boldsymbol{y}$ belongs to $S_1 \cap \cdots \cap S_m$.

Since $\boldsymbol{x}, \boldsymbol{y} \in S_1 \cap \cdots \cap S_m$, certainly it's true that $\boldsymbol{x} \in S_1$ and $\boldsymbol{y} \in S_1$. Then since S_1 is a convex set by assumption, $\boldsymbol{v} \in S_1$.

Similarly, for each $i = 1, \dots, m$, $\boldsymbol{v} \in S_i$. Therefore $\boldsymbol{v} \in S_1 \cap \cdots \cap S_m$. Q.E.D.

1.3 #34. Proof. Suppose that $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ both satisfy $A\boldsymbol{x} > b$. We must show that any point $\boldsymbol{v}$ on the line segment joining $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ also satisfies $A\boldsymbol{x} > b$.

If $\boldsymbol{v} = \boldsymbol{x}_1$ or $\boldsymbol{x}_2$, the inequality is satisfied by assumption. Any other $\boldsymbol{v}$ on the line segment in question has the form $\boldsymbol{v} = \lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$ for some real number λ such that $0 < \lambda < 1$. In particular, $\lambda > 0$ and $1 - \lambda > 0$.

We know that $A\boldsymbol{x}_1 > b$ and $A\boldsymbol{x}_2 > b$. Multiply these by λ and $1-\lambda$ respectively, and add. This gives the inequality below. The rest basic matrix algebra:

$$A\boldsymbol{v} = A(\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) = \lambda(A\boldsymbol{x}_1) + (1-\lambda)(A\boldsymbol{x}_2) > \lambda b + (1-\lambda)b = b.$$

So $A\boldsymbol{v} > b$. Q.E.D.

1.4, #16. By definition (it's in Section 1.4), $\boldsymbol{x}$ is a convex combination of points $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m \in \boldsymbol{R}^n$ if and only if

$$\boldsymbol{x} = a_1\boldsymbol{x}_1 + \cdots + a_m\boldsymbol{x}_m \qquad \text{for some scalars } a_1, \dots, a_m \text{ such that}$$
$$a_i \geq 0 \text{ for all } i = 1, \dots, m, \text{ and } a_1 + \cdots + a_m = 1.$$

Now suppose that objective function of a LPP is z, and the optimal value z_{opt} is attained at $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$. Then, $z(\boldsymbol{x}_1) = z(\boldsymbol{x}_2) = \cdots = z(\boldsymbol{x}_m) = z_{opt}$. Let $\boldsymbol{x} = a_1\boldsymbol{x}_1 + \cdots + a_m\boldsymbol{x}_m$ be any convex combination as above. We must show that $z(\boldsymbol{x}) = z_{opt}$.

Let $\boldsymbol{c}$ be the cost vector for the LPP. Thus, $z = \boldsymbol{c}^T\boldsymbol{x}$. Using basic properties of matrix operations,

$$\begin{aligned} z(\boldsymbol{x}) &= \boldsymbol{c}^T\boldsymbol{x} = \boldsymbol{c}^T(a_1\boldsymbol{x}_1 + \cdots + a_m\boldsymbol{x}_m) = \boldsymbol{c}^T(a_1\boldsymbol{x}_1) + \cdots + \boldsymbol{c}^T(a_m\boldsymbol{x}_m) \\ &= a_1(\boldsymbol{c}^T\boldsymbol{x}_1) + \cdots + a_m(\boldsymbol{c}^T\boldsymbol{x}_m) = a_1 z(\boldsymbol{x}_1) + \cdots + a_m z(\boldsymbol{x}_m) \\ &= a_1 z_{opt} + \cdots + a_m z_{opt} = (a_1 + \cdots + a_m) z_{opt} = 1 \cdot z_{opt} \\ &= z_{opt}. \qquad \text{Q.E.D.} \end{aligned}$$

(Two related side points: a) It is also true that if $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ are all feasible, then $\boldsymbol{x}$ is feasible, but the problem is not asking for a proof of this. b) The property $a_1 \geq 0, \dots, a_m \geq 0$ never needs to be used in the given problem. But it is necessary for showing that $\boldsymbol{x}$ is feasible.)

A. First let's prove it for vectors in $\boldsymbol{R}^1$ by showing that $y = -\log x$ is a convex function, that is, for all positive real numbers x_1, x_2 and all λ such that $0 \leq \lambda \leq 1$, it is true that

$$(1) \qquad -\log(\lambda x_1 + (1-\lambda)x_2) \leq \lambda(-\log x_1) + (1-\lambda)(-\log x_2).$$

One way to show this is to use the geometric interpretation of (1), that the graph of $y = -\log x$, for $x_1 \leq x \leq x_2$, lies below the line segment joining the points $(x_1, \log x_1)$ and $(x_2, \log x_2)$ on the graph. This is true because the graph is concave up, which in turn is true because $y'' = (-1/x)' = 1/x^2 > 0$ for all x. **Note:** This explanation would have been adequate.

Another way, which is more analytical and is rigorous, is essentially to prove that any real-valued function defined on any interval in $\boldsymbol{R}$, and with a positive second derivative everywhere on that interval, is convex. Write

$f(x) = -\log x$ and use Taylor's formula with remainder, with $n = 1$, "centered" at $x = x_0$:

$$f(x) = f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(c)(x-x_0)^2 \text{ for some } c \text{ between } x \text{ \& } x_0.$$

This is true for all positive x and x_0. Since $f''(c) > 0$ for all c between x_1 and x_2,

$$f(x) > f(x_0) + f'(x_0)(x - x_0).$$

Now, a) substitute $x = x_1$ and multiply through by λ (which is ≥ 0 by assumption); b) then do the same with $x = x_2$ ad $1 - \lambda$; and c) add the results:

$$\lambda f(x_1) \geq \lambda f(x_0) + \lambda f'(x_0)(x_1 - x_0)$$

$$(1-\lambda)f(x_2) \geq (1-\lambda)f(x_0) + (1-\lambda)f'(x_0)(x_2 - x_0)$$

$$\begin{aligned} \lambda f(x_1) + (1-\lambda)f(x_2) &\geq \lambda f(x_0) + \lambda f'(x_0)(x_1 - x_0) \\ &\quad + (1-\lambda)f(x_0) + (1-\lambda)f'(x_0)(x_2 - x_0) \\ &= f(x_0) + f'(x_0)\big[\lambda(x_1 - x_0) + (1-\lambda)(x_2 - x_0)\big]. \\ &= f(x_0) + f'(x_0)\big[\lambda x_1 + (1-\lambda)x_2 - x_0\big]. \end{aligned}$$

Substitute $x_0 = \lambda x_1 + (1-\lambda)x_2$ and the second term on the right becomes 0, and the inequality is what we wanted to prove.

Now what about $\boldsymbol{R}^n$? Let $\boldsymbol{x}_1 = [x_{11} \ldots, x_{1n}]^T$ and $\boldsymbol{x}_2 = [x_{21} \ldots, x_{2n}]^T$ be vectors in $\boldsymbol{R}^n$, with all coordinates positive (so that the logs will be

defined). Then

$$\begin{aligned} f(\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) &= f\left(\lambda\begin{bmatrix} x_{11} \\ \vdots \\ x_{1n}\end{bmatrix} + (1-\lambda)\begin{bmatrix} x_{21} \\ \vdots \\ x_{2n}\end{bmatrix}\right) \\ &= f\left(\begin{bmatrix} \lambda x_{11} + (1-\lambda)x_{21} \\ \vdots \\ \lambda x_{1n} + (1-\lambda)x_{2n}\end{bmatrix}\right) \\ &= -\sum_{i=1}^{n}\log(\lambda x_{1i} + (1-\lambda)x_{2i}) \\ &\leq -\sum_{i=1}^{n}\big[\lambda\log x_{1i} + (1-\lambda)\log x_{2i}\big] \text{ (we've proved it for } \boldsymbol{R}^1) \\ &= -\lambda\sum_{i=1}^{n}\log x_{1i} \; - \; (1-\lambda)\sum_{i=1}^{n}\log x_{2i} \\ &= \lambda f(\boldsymbol{x}_1) + (1-\lambda)f(\boldsymbol{x}_2). \qquad \text{Q.E.D.} \end{aligned}$$

B. (Compare #34 above) Proof. Suppose that $f(\boldsymbol{x}) \leq b$ is satisfied by $\boldsymbol{x} = \boldsymbol{x}_1$ and $\boldsymbol{x} = \boldsymbol{x}_2$. We must show that it is satisfied by any $\boldsymbol{x}$ on the line segment joining $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. Any such $\boldsymbol{x}$ has the form $\boldsymbol{x} = \lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$ for some λ such that $0 \leq \lambda \leq 1$.

We know that

$$f(\boldsymbol{x}_1) \leq b \quad \text{and} \quad f(\boldsymbol{x}_2) \leq b.$$

Multiply these by λ and $1-\lambda$ respectively. Since $\lambda \geq 0$ and $1-\lambda \geq 0$, the inequalities remain true:

$$\lambda f(\boldsymbol{x}_1) \leq \lambda b \quad \text{and} \quad (1-\lambda)f(\boldsymbol{x}_2) \leq (1-\lambda)b.$$

Add these (the second inequality below) and use the fact that f is convex (the first inequality below):

$$f(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1)+(1-\lambda)f(\boldsymbol{x}_2) \leq \lambda b+(1-\lambda)b = b. \qquad \text{Q.E.D.}$$