Exercise 3.18 [Boyd & Vandenberghe, 2004]

Adapt the proof of concavity of the log-determinant function in §3.1.5 to show the following.
(a) $f(X) = \mathrm{tr}(X^{-1})$ is convex on $\mathbf{dom}\, f = \mathbf{S}^n_{++}$.
(b) $f(X) = (\det X)^{1/n}$ is concave on $\mathbf{dom}\, f = \mathbf{S}^n_{++}$.

Ans: Here only does part (a).
(a) It suffices to show $g(t) = \mathrm{tr}((X + tZ)^{-1})$ is convex w.r.t. $t \in [t_0, t_1]$ where $X \in \mathbf{S}^n_{++}$, $Z \in \mathbf{S}^n$ and $[t_0, t_1]$ is an interval such that $X + tZ$ remains positive definite. Use this property of positive definite matrix, write

$$X = S^{-1}S^{-T}, Z = S^{-1}LS^{-T}$$

then

$$\begin{aligned}
g(t) &= \mathrm{tr}((X + tZ)^{-1}) \\
&= \mathrm{tr}((S^{-1}(I + tL)S^{-T})^{-1}) \\
&= \mathrm{tr}(S(I + tL)^{-1}S^T) \\
&= \mathrm{tr}(S^T S(I + tL)^{-1}) \\
g'(t) &= \mathrm{tr}\left(S^T S \frac{d}{dt}(I + tL)^{-1}\right) \\
&= -\mathrm{tr}(S^T S(I + tL)^{-1}L(I + tL)^{-1})
\end{aligned}$$

Here I made use of the identity

$$\begin{aligned}
I &= A^{-1}(t)A(t) \\
O &= \left(\frac{d}{dt}A^{-1}(t)\right)A(t) + A^{-1}(t)A'(t) \\
\frac{d}{dt}A^{-1}(t) &= -A^{-1}(t)A'(t)A^{-1}(t)
\end{aligned}$$

Now that $I + tL$ and L are both diagonal, they commute.

$$\begin{aligned}
g'(t) &= -\mathrm{tr}(S^T S(I + tL)^{-1}L(I + tL)^{-1}) \\
&= -\mathrm{tr}(S^T SL(I + tL)^{-2}) \\
g''(t) &= \mathrm{tr}(S^T SL((I + tL)^{-2}L(I + tL)^{-1} + (I + tL)^{-1}L(I + tL)^{-2})) \\
&= 2\mathrm{tr}(SL^2(I + tL)^{-3}S^T)
\end{aligned}$$

Since t is such that $X + tZ = S^{-1}(I + tL)S^{-T} \succ 0$, $I + tL$ must be positive diagonal matrix, and $(I + tL)^{-3} \succ 0$; alongside with $L^2 \succeq 0$, we see $L^2(I + tL)^{-3} \succeq 0$, therefore

$$SL^2(I + tL)^{-3}S^T \succeq 0$$

and $g''(t) \geq 0$ (*i.e.* g is convex). $\square$

Exercise 3.19 [Boyd & Vandenberghe, 2004]

Nonnegative weighted sums and integrals.
(a) Show that $f(x) = \sum_{i=1}^{r} \alpha_i x_{[i]}$ is a convex function of x, where $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_r \geq 0$, and $x_{[i]}$ denotes the ith largest component of x. (You can use the fact that $f(x) = \sum_{i=1}^{k} x_{[i]}$ is convex on $\mathbb{R}^n$.)
(b) Let $T(x, \omega)$ denote the trigonometric polynomial

$$T(x, \omega) = x_1 + x_2 \cos \omega + x_3 \cos 2\omega + \cdots + x_n \cos(n-1)\omega.$$

Show that the function

$$f(x) = -\int_0^{2x} \log T(x, \omega) d\omega$$

is convex on $\{x \in \mathbb{R}^n \mid T(x, \omega) > 0, 0 \leq \omega \leq 2\pi\}$.

Ans: Here only does part (a).
(a) Let $\beta_i = \alpha_i - \alpha_{i+1}$; then observe

$$\begin{aligned}
f(x) &= \sum_{i=1}^{r} \alpha_i x_{[i]} \\
&= \alpha_r \sum_{i=1}^{r} x_{[i]} + (\alpha_{r-1} - \alpha_r) \sum_{i=1}^{r-1} x_{[i]} + \cdots + (\alpha_1 - \alpha_2) \sum_{i=1}^{1} x_{[i]} \\
&= \alpha_r \sum_{i=1}^{r} x_{[i]} + \beta_{r-1} \sum_{i=1}^{r-1} x_{[i]} + \cdots + \beta_1 \sum_{i=1}^{1} x_{[i]}
\end{aligned}$$

is a linear combination of convex functions with nonnegative coefficients and therefore convex. $\square$

Exercise 3.22 [Boyd & Vandenberghe, 2004]

Composition rules. Show that the following functions are convex.
(a) $f(x) = -\log(-\log(\sum_{i=1}^m e^{a_i^T x+b_i}))$ on $\mathbf{dom}\, f = \{x \mid \sum_{i=1}^m e^{a_i^T x+b_i} < 1\}$. You can use the fact that $\log(\sum_{i=1}^n e^{y_i})$ is convex.
(b) $f(x,u,v) = -\sqrt{uv - x^T x}$ on $\mathbf{dom}\, f = \{(x,u,v) \mid uv > x^T x, u, v > 0\}$. Use the fact that $x^T x/u$ is convex in (x,u) for $u > 0$, and that $-\sqrt{x_1 x_2}$ is convex on $\mathbb{R}^2_{++}$.
(c) $f(x,u,v) = -\log(uv - x^T x)$ on $\mathbf{dom}\, f = \{(x,u,v) \mid uv > x^T x, u, v > 0\}$.
(d) $f(x,t) = -(t^p - \|x\|_p^p)^{1/p}$ where $p > 1$ and $\mathbf{dom}\, f = \{(x,t) \mid t \geq \|x\|_p\}$. You can use the fact that $\|x\|_p^p/u^{p-1}$ is convex in (x,u) for $u > 0$ (see exercise 3.23), and that $-x^{1/p}y^{1-1/p}$ is convex on $\mathbb{R}^2_+$ (see exercise 3.16).
(e) $f(x,t) = -\log(t^p - \|x\|_p^p)$ where $p > 1$ and $\mathbf{dom}\, f = \{(x,t) \mid t > \|x\|_p\}$. You can use the fact that $\|x\|_p^p/u^{p-1}$ is convex in (x,u) for $u > 0$ (see exercise 3.23).

Ans: Here only does part (c).
(c) Let $g(x,u,v) = uv - x^T x$, then $\nabla g(x,u,v) = [-2x^T, v, u]$ and

$$H_g(x,u,v) = \begin{bmatrix} -2I & & \\ & & 1 \\ & 1 & \end{bmatrix}$$

and for any $(x,u,v) \in \mathbf{dom}\, f$,

$$[x^T, u, v] \begin{bmatrix} -2I & & \\ & & 1 \\ & 1 & \end{bmatrix} \begin{bmatrix} x \\ u \\ v \end{bmatrix} = -2x^T x + 2uv > 0$$

That is, g is convex on $\mathbf{dom}\, f$. Now that $h = -\log(\cdot)$ is convex and nonincreasing, we conclude $f = -\log(g(\cdot))$ is convex on $\mathbf{dom}\, f$ as well. $\square$

Additional Exercise 2.5 [Boyd & Vandenberghe, 2017]

A perspective composition rule [Marèchal]. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function with $f(0) \leq 0$.
(a) Show that the perspective $tf(x/t)$, with domain $\{(x,t) \mid t > 0, x/t \in \textbf{dom}\, f\}$, is nonincreasing as a function of t.
(b) Let g be concave and positive on its domain. Show that the function

$$h(x) = g(x)f(x/g(x)), \quad \textbf{dom}\, h = \{x \in \textbf{dom}\, g \mid x/g(x) \in \textbf{dom}\, f\}$$

is convex.
(c) As an example, show that

$$h(x) = \frac{x^T x}{\left(\prod_{k=1}^{n} x_k\right)^{1/n}}, \quad \textbf{dom}\, h = \mathbb{R}^n_{++}$$

is convex.

Ans: Here only does part (a) and (b).
(a) Fix $x \in \mathbb{R}^n$, for any $0 < t_1 < t_2$ such that $x/t_1, x/t_2 \in \textbf{dom}\, f$, note that

$$\frac{x}{t_2} = \frac{t_1}{t_2}\frac{x}{t_1} + \frac{t_2 - t_1}{t_2} \cdot 0$$

By convexity of f,

$$\begin{aligned} f(x/t_2) &\leq \frac{t_1}{t_2} f(x/t_1) + \left(1 - \frac{t_1}{t_2}\right) f(0) \\ t_2 f(x/t_2) - t_1 f(x/t_1) &\leq \left(1 - \frac{t_1}{t_2}\right) f(0) \leq 0 \end{aligned}$$

we conclude that $t_2 f(x/t_2) \leq t_1 f(x/t_1)$, and f is nonincreasing.

(b) First I want to show for $\alpha, \beta \in [0,1]$ with $\alpha + \beta \leq 1$ and $x, y \in \textbf{dom}\, f$,

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y)$$

Let $t = \alpha + \beta \leq 1$ and $\theta = \alpha/t$, then $1 - \theta = \beta/t$. The original inequality now becomes

$$f(\theta t x + (1-\theta) t y) \leq \theta t f(x) + (1-\theta) t f(y)$$

This can be proven since f is convex and the perspective $h(t,x) = tf(x/t)$ is nonincreasing w.r.t. t,

$$\begin{aligned} f(\theta t x + (1-\theta) t y) &\leq \theta f(tx) + (1-\theta) f(ty) \\ &= \theta h(1, tx) + (1-\theta) h(1, ty) \\ &\leq \theta h(t, tx) + (1-\theta) h(t, ty) \\ &= \theta t f(tx/t) + (1-\theta) t f(ty/t) = \theta t f(x) + (1-\theta) t f(y) \end{aligned}$$

Now for general $\theta \in [0,1]$ and $x, y \in \mathbf{dom}\, h$, since g is concave,

$$\alpha = \frac{\theta g(x)}{g(\theta x + (1-\theta)y)}, \beta = \frac{(1-\theta)g(y)}{g(\theta x + (1-\theta)y)} \in [0,1]$$

and $\alpha + \beta = \frac{\theta g(x) + (1-\theta)g(y)}{g(\theta x + (1-\theta)y)} \leq 1$. Apply the above "lemma",

$$\begin{aligned}
h(\theta x + (1-\theta)y) &= g(\theta x + (1-\theta)y) f\left(\frac{\theta x + (1-\theta)y}{g(\theta x + (1-\theta)y)}\right) \\
&= g(\theta x + (1-\theta)y) f\left(\frac{\theta g(x)}{g(\theta x + (1-\theta)y)}\frac{x}{g(x)} + \frac{(1-\theta)g(y)}{g(\theta x + (1-\theta)y)}\frac{y}{g(y)}\right) \\
&= g(\theta x + (1-\theta)y) f\left(\alpha\frac{x}{g(x)} + \beta\frac{y}{g(y)}\right) \\
&\leq g(\theta x + (1-\theta)y)\left(\alpha f\left(\frac{x}{g(x)}\right) + \beta f\left(\frac{y}{g(y)}\right)\right) \\
&= \theta g(x) f\left(\frac{x}{g(x)}\right) + (1-\theta)g(y) f\left(\frac{y}{g(y)}\right) \\
&= \theta h(x) + (1-\theta)h(y)
\end{aligned}$$

(c) Consider $f(x) = x^T x$ and $g(x) = \left(\prod_{k=1}^{n} x_k\right)^{1/n}$; they are both common convex functions. Then

$$\begin{aligned}
h(x) &= g(x) f(x/g(x)) \\
&= \left(\prod_{k=1}^{n} x_k\right)^{1/n} \frac{x^T x}{\left(\prod_{k=1}^{n} x_k\right)^{2/n}} \\
&= \frac{x^T x}{\left(\prod_{k=1}^{n} x_k\right)^{1/n}}
\end{aligned}$$

is convex, too, by part (b). □

Additional Exercise 2.30 [Boyd & Vandenberghe, 2017]

Huber penalty. The infimal convolution of two functions f and g on $\mathbb{R}^n$ is defined as

$$h(x) = \inf_y (f(y) + g(x-y))$$

(see exercise 2.17). Show that the infimal convolution of $f(x) = \|x\|_1$ and $g(x) = (1/2)\|x\|_2^2$, *i.e.* , the function

$$h(x) = \inf_y (f(y) + g(x-y)) = \inf_y \left(\|y\|_1 + \frac{1}{2}\|x-y\|_2^2 \right)$$

is the *Huber penalty*

$$h(x) = \sum_{i=1}^n \phi(x_i), \quad \phi(u) = \begin{cases} u^2/2, |u| \le 1 \\ |u| - 1/2, |u| > 1 \end{cases}$$

Ans: Observe first

$$\begin{aligned} h(x) &= \inf_{y \in \mathbb{R}^n} \left(\sum_{i=1}^n |y_i| + \frac{1}{2}(x_i - y_i)^2 \right) \\ &= \sum_{i=1}^n \inf_{y_i} \left(|y_i| + \frac{1}{2}(x_i - y_i)^2 \right) \end{aligned}$$

Let $g(x_i) = \inf_{y_i} |y_i| + \frac{1}{2}(x_i - y_i)^2$. In order to show $g(x_i) = \phi(x_i)$, observe first that for $x_i \ge 0$, the minimizing y_i must also be nonnegative and vice versa. Consider the case $x_i \ge 0$, then

$$g(x_i) = \inf_y y + \frac{1}{2}(x_i - y)^2$$

And the minimizer y satisfies

$$\frac{d}{dy}\left(y + \frac{1}{2}(x_i - y)^2 \right) = 1 - (x_i - y) = 0$$

$$\frac{d^2}{dy^2}\left(y + \frac{1}{2}(x_i - y)^2 \right) = 1 > 0$$

therefore $\bar{y} = x_i - 1$. If $x_i \ge 1$ then the minimizer will be in the scope (we assumed we are minimizing among all $y \ge 0$) and

$$g(x_i) = \bar{y} + \frac{1}{2}(x_i - \bar{y})^2 = x_i - \frac{1}{2}$$

However, if $x_i < 1$, then 0 would be a better minimizer and

$$g(x_i) = 0 + \frac{1}{2}(x_i - 0)^2 = \frac{1}{2}x_i^2$$

Concluding above, we have

$$g(x_i) = \begin{cases} x_i - 1/2 & ,1 \leq x_i \\ x_i^2/2 & ,0 < x_1 < 1 \end{cases}$$

Since the same should hold analogously for $x_i < 0$, we prove $g(x_i) = \phi(x_i)$ for x_i in all ranges. $\square$

Additional Exercise 2.31 [Boyd & Vandenberghe, 2017]

Suppose the function $h : \mathbb{R} \to \mathbb{R}$ is convex, nondecreasing, with $\mathbf{dom}\, h = \mathbb{R}$, and $h(t) = h(0)$ for $t \le 0$.
(a) Show that the function $f(x) = h(\|x\|_2)$ is convex on $\mathbb{R}^n$.
(b) Show that the conjugate of f is $f^*(y) = h^*(\|y\|_2)$.
(c) As an example, derive the conjugate of $f(x) = (1/p)\|x\|_2^p$ for $p > 1$, by applying the result of part (b) with the function

$$h(t) = \frac{1}{p}\max\{0, t\}^p = \begin{cases} \frac{1}{p}t^p, t \ge 0 \\ 0, t < 0 \end{cases}$$

Ans:
(a) Let $g(x) = \|x\|_2$, then

$$\begin{aligned}
\nabla g(x) &= \frac{x^T}{\|x\|_2} \\
H_g(x) &= \frac{1}{\|x\|_2^2}\left(I\|x\|_2 - \frac{x}{\|x\|_2}x^T\right) \\
&= \frac{1}{\|x\|_2}\left(I - \frac{1}{\|x\|_2^2}xx^T\right) \\
v^T H_g(x) v &= \frac{1}{\|x\|_2}\left(\|v\|_2 - \frac{(x^T v)^2}{\|x\|_2^2}\right) \ge 0 \text{ (by Cauchy inequality)}
\end{aligned}$$

Since h is convex and nondecreasing and $g = \|\cdot\|_2$ is convex, the composition $f = h \circ g$ must be convex, too.

(b) By definition,

$$\begin{aligned}
f^*(y) &= \sup_{x\in\mathbb{R}^n} x^T y - f(x) \\
&= \sup_{x\in\mathbb{R}^n} x^T y - h(\|x\|_2) \\
&= \max\left\{\sup_{x\neq 0} \|x\|_2 \left(\frac{x}{\|x\|_2}\right)^T y - h(\|x\|_2), -h(0)\right\} \\
&= \sup_{x\neq 0} \|x\|_2 \left(\frac{x}{\|x\|_2}\right)^T y - h(\|x\|_2) \qquad \text{(because } -h \text{ is nonincreasing)} \\
&= \sup_{x\neq 0} \|x\|_2\|y\|_2 - h(\|x\|_2) \\
&= \sup_{t>0} t\|y\|_2 - h(t)
\end{aligned}$$

The last equality follows from that the first term is maximized when x is parallel to y. Now observe the definition of h^*

$$h^*(\|y\|_2) = \sup_{t\in\mathbb{R}} t\|y\|_2 - h(t)$$

This differs from $f^*(y)$ only by the range of the minimizing variable (it's $t > 0$ for f^* and $t \in \mathbb{R}$ for h^*); however, since $h(t) = h(0)$ for $t \leq 0$, that is, for any $t < 0$,

$$t\|y\|_2 - h(t) = t\|y\|_2 - h(0) \leq 0 \cdot \|y\|_2 - h(0)$$

The minimizer for h^* can't fall in $(-\infty, 0)$. Therefore we conclude $f^*(y) = h^*(\|y\|_2)$.

(c) From the problem settings, $f(x) = \frac{1}{p}\|x\|_2^p = h(\|x\|_2)$. To show f is convex, it suffices to show that h is convex, nondecreasing and $h(t) = h(0)$ for $t \leq 0$. h is convex since

$$h'(t) = \begin{cases} t^{p-1}, t \geq 0 \\ 0, t < 0 \end{cases}$$

is continuous and nondecreasing; h is nondecreasing follows from $p > 1$ and $h(t) = h(0)$ for $t \leq 0$ is just a part of the definition. $\square$

Additional Exercise 3.17 [Boyd & Vandenberghe, 2017]

Minimum fuel optimal control. Solve the minimum fuel optimal control problem described in exercise 4.16 of *Convex Optimization*, for the instance with problem data

$$A = \begin{bmatrix} -1 & 0.4 & 0.8 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 0 \\ 0.3 \end{bmatrix}, x_{des} = \begin{bmatrix} 7 \\ 2 \\ -6 \end{bmatrix}, N = 30.$$

You can do this by forming the LP you found in your solution of exercise 4.16, or more directly using CVX. Plot the actuator signal $u(t)$ as a function of time t.

Exercise 4.16 [Boyd & Vandenberghe, 2004]

Minimum fuel optimal control. We consider a linear dynamical system with state $x(t) \in \mathbb{R}^n$, $t = 0, \cdots, N$, and actuator or input signal $u(t) \in \mathbb{R}$, for $t = 0, \cdots, N-1$. The dynamics of the system is given by the linear recurrence

$$x(t+1) = Ax(t) + bu(t), t = 0, \cdots, N-1,$$

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 $u(0), \cdots, u(N-1)$ so as to minimize the total fuel consumed, which is given by

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

subject to the constraint that $x(N) = x_{des}$, where N is the (given) time horizon, and $x_{des} \in \mathbb{R}^n$ 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

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

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.

Ans: The constraint is linear in $u(t)$, since $x(N)$ is a linear combination of $A^{N-1}b, \cdots, Ab, b$

with coefficients $u(0), \cdots, u(N-2), u(N-1)$,

$$\begin{aligned}
x(N) &= Ax(N-1) + bu(N-1) \\
&= A^2x(N-2) + Abu(N-2) + bu(N-1) \\
&= \cdots \\
&= A^N x(0) + A^{N-1}bu(0) + A^{N-2}bu(1) + \cdots + Abu(N-2) + bu(N-1) \\
&= A^{N-1}bu(0) + A^{N-2}bu(1) + \cdots + Abu(N-2) + bu(N-1)
\end{aligned}$$

and the original constraint is not much but

$$x_{des} = A^{N-1}bu(0) + A^{N-2}bu(1) + \cdots + Abu(N-2) + bu(N-1)$$

On the other hand, by introducing new variable[1] $y(t)$ with constraints

$$\pm u(t), \pm 2u(t) - 1 \leq y(t)$$

we find

$$\begin{aligned}
f(u(t)) &= \begin{cases} |u(t)|, |u(t)| \leq 1 \\ 2|u(t)| - 1, |u(t)| > 1 \end{cases} \\
&= \max\{\pm u(t), \pm 2u(t) - 1\} \\
&= y(t)
\end{aligned}$$

That is, the problem comes with $4 * N + 3$ constraints on variables $u(t)$ and $y(t)$ with objective function $F = \sum_{i=0}^{N-1} y(t)$. Or more formally speaking, the problem is

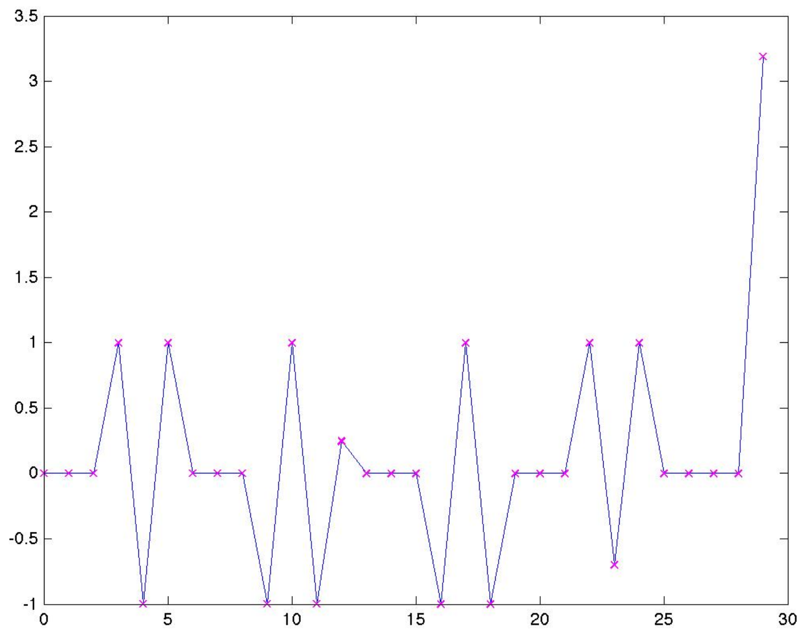
$$\begin{array}{rl}
\text{minimize} & \mathbf{1}y \\
\text{subject to} & x_{des} = \mathcal{K}'(A,b)u \\
& u \leq y \\
& -u \leq y \\
& 2u - \mathbf{1} \leq y \\
& -2u - \mathbf{1} \leq y
\end{array}$$

where $\mathcal{K}'(A,b)$ is the Krylov matrix in reverse order

$$\mathcal{K}'(A,b) = \begin{bmatrix} A^{N-1}b & \cdots & Ab & b \end{bmatrix}$$

Here's the graph of $u(t)$ as a function of t and the codes used to generate the solution. I implemented the direct problem using CVX (plotted as blue line) as well as the linear programming problem (plotted as magenta crosses); it's encouraging that they give the same solution!

[1] credit to Minh Pham for his great hint!




```
f=@(x) max(abs(x), 2*abs(x)-1);
F=@(u) sum(f(u));
A = [-1,0.4,0.8;1,0,0;0,1,0];
b = [1;0;0.3];
x_des = [7;2;-6];
N = 30;
KAb = zeros(3,30);
KAb(:,30) = b;
for i=29:-1:1
    KAb(:,i) = A * KAb(:, i+1);
end

cvx_begin
    variable u(N);
    minimize F(u);
    subject to
        KAb*u == x_des;
cvx_end
plot(0:29,u','-b');

clear u
%% alternative solution (LP)
cvx_begin
    variable u(2*N);
```

```
    minimize ones(1,N)*u(N+1:end);
    subject to
        KAb*u(1:N) == x_des;
        u(1:N) <= u(N+1:end);
        -u(1:N) <= u(N+1:end);
        2*u(1:N)-ones(N,1) <= u(N+1:end);
        -2*u(1:N)-ones(N,1) <= u(N+1:end);
cvx_end
hold on
handle = plot(0:29,u(1:N)','mx','LineWidth',1.2);
saveas(handle, 'hw3_p317','jpg');
```