EE 236B HW3
Zhiyuan Cao
304397496
02/02/2017

T3.18

(a)

Define $g(t) = f(Z + tV)$ and restrict g to the interval of values of t for which $Z + tV \succ 0$. We have

$$\begin{aligned} g(t) &= \mathbf{tr}((Z + tV)^{-1}) \\ &= \mathbf{tr}((Z^{1/2}(I + tZ^{-1/2}VZ^{-1/2})Z^{1/2})^{-1}) \\ &= \mathbf{tr}(Z^{-1}(I + tZ^{-1/2}VZ^{-1/2})^{-1}) \\ &= \mathbf{tr}(Z^{-1}(I + tQ\Lambda Q^T)^{-1}) \\ &= \mathbf{tr}(Q^T Z^{-1} Q(I + t\Lambda)^{-1}) \\ &= \sum_{i=1}^{n} (Q^T Z^{-1} Q)_{ii}(1 + t\lambda_i)^{-1} \end{aligned}$$

Where we had $Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^T$. $g(t)$ can be seen as a positive weighted sum of convex functions $1/(1 + t\lambda_i)^{-1}$ and therefore is convex.

T3.19

(a)

We introduce $y \in R^n$, where

$$y_{[i]} = \begin{cases} \alpha_i x_{[i]} & i \leq r \\ \alpha_r x_{[i]} & \text{otherwise.} \end{cases}$$

With $\alpha_1 \geq \alpha_2 \geq ... \geq \alpha_r \geq 0$, we have $y_{[1]} \geq y_{[2]} \geq ... \geq y_{[n]}$ and the original weighted sum problem can be written as

$$f(y) = \sum_{i=1}^{r} y_{[i]}$$

Therefore it is convex.

T3.22

(c)

$$f(x) = -\log(u(v - \frac{x^T x}{u})) = -\log u - \log(v - \frac{x^T x}{u})$$

From composition rule, with $v - \frac{x^T x}{u}$ being concave and $-\log x$ concave and nondecreasing, $-\log(v - \frac{x^T x}{u})$ is concave. Also $-\log u$ is concave and therefore $f(x)$ is the sum of two concave function and therefore is also concave.

A2.5

(a) from $g(x,t) = tf(x/t)$ we have

$$\begin{aligned}\frac{\partial g}{\partial t} &= f(x/t) - \frac{x}{t} f(x/t) \\ &= f(x/t) + (0 - \frac{x}{t}) f(x/t) \leq f(0) \\ &\leq 0\end{aligned}$$

This is followed by first-order condition. Therefore $g(x,t)$ is nonincreasing as a function of t.

(b)

We can write $h(x) = f(g(x,t))$, where $f(x,t) = tf(x/t)$ and $g(x,t) = g(x)$. From composition rule $g(x,t)$ is concave and $f(x,t)$ is nonincreasing and convex, hence $h(x)$ is convex.

A2.30

We can rewrite $h(x)$ as

$$h(x) = \inf_y (\sum_{i=1}^{n} |y_i| + \frac{1}{2} \sum_{i=1}^{n} (x_i - y_i)^2)$$

To minimize $h(x)$ over y it is equivalent to minimizing sum of each term, call it $g(x_i, y_i)$, over y_i separately

$$\sum_{i=1}^{n} \inf_{y_i} (|y_i| + \frac{1}{2}(x_i - y_i)^2) = \sum_{i=1}^{n} g(x_i, y_i)$$

To see this is the *Huber penalty*, we set

$$\frac{\partial g(x_i, y_i)}{\partial y_i} = 0 = \begin{cases} y_i - x_i + 1 & y_i \geq 0 \\ y_i - x_i - 1 & \text{otherwise.} \end{cases}$$

and get the following:

$$y_i = \begin{cases} x_i - 1 & y_i \geq 0, x_i \geq 1 \\ x_i + 1 & y_i < 0, x_i \leq -1 \\ \text{no solution} & \text{otherwise} \end{cases}$$

For the first two cases, combining them and we acquire

$$\phi(u) = \{(|u| - 1/2) \mid |u| \geq 1\}$$

For the case where $|x_i| \leq 1$, when $y_i \geq 0$ we have $\partial g(x_i, y_i)/\partial y_i \geq 0$ and therefore is nondecreasing over y_i. Hence $\inf_{y_i} g(x_i, y_i) = 1/2(x_i)^2$ with y_i set to 0. Same argument can be made in the case where $y_i < 0$. Combining all cases we get the *Huber penalty*, i.e.,

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

A2.31

(a)

Using the composition rule, $f(x) = h(g(x))$ where $g(x) = ||x||_2$, since $g(x)$ is convex on R^n and $h(x)$ is nondecreasing and convex, $f(x)$ is convex.

(b)

$$\begin{aligned} f^*(y) &= \sup_x (y^T x - h(||x||_2)) \\ &= \sup_{t \geq 0} \sup_{||x||_2 = t} (y^T x - h(t)) \\ &\leq \sup_{t \geq 0} \sup_{||x||_2 = t} (||y||_2 ||x||_2 - h(t)) \\ &= \sup_{t \geq 0} (t||y||_2 - h(t)) \\ &= h^*(||y||_2) \end{aligned}$$

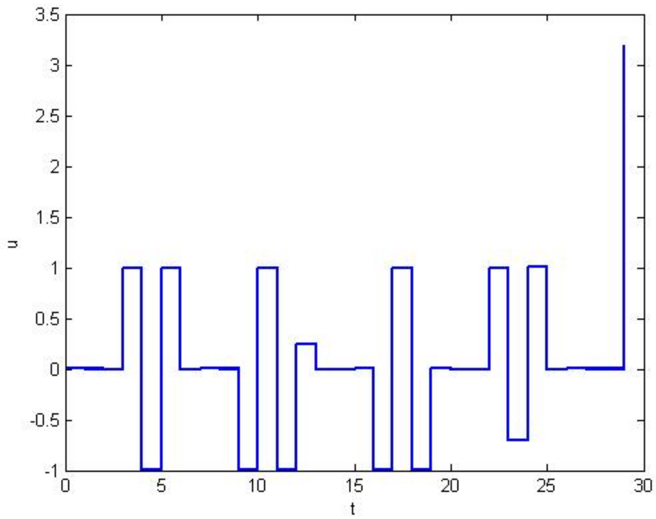

Figure 1: Output $u(t)$

(c)

Since$||y||_2 \geq 0$, we have

$$h^*(||y||_2) = \sup_t (t||y||_2 - h(t))$$
$$= t||y||_2 - pt^p$$

It reaches maximum at $||y||_2 = t^{p-1}$ and

$$f^*(y) = h^*(||y||_2) = \frac{p-1}{p}||y||_2^{\frac{p}{p-1}}, \qquad \forall y \in R^n$$

A3.17

See Figure 1 and M-code

```
clc;clear;
N = 30;n = 3;
A = [-1, .4, .8; 1, 0, 0; 0, 1, 0];
b = [1; 0; 0.3];
x_des = [7; 2; -6];
x_init = zeros(n, 1);
cvx_begin
    variable x(n, N+1)
    variable u(1, N)
    minimize (sum(max(abs(u), 2*abs(u)-1)))
    subject to
        x(:, 2 : N+1) == A*x(:, 1 : N) + b*u
```

```
        x(:, 1) == x_init;
        x(:, N+1) == x_des;
cvx_end
stairs(0 : N-1, u, 'LineWidth', 2)
xlabel('t')
ylabel('u')
```