ECE 236B - HW 2

Peter Raciappo
103953689

$\boxed{3.18}$

Show
(a.) $f(X) = tr(X^{-1})$ is convex on $dom\ f = S^n_{++}$.

Let $X = Z + tV$, $Z, V \in S^n$ with $Z \succ 0$, $t \in \mathbb{R}_+$, and let $g(t) = f(Z + tV)$. We have:

$$g(t) = tr[(Z + tV)^{-1}].$$

Now, $(I + tZ^{-\frac{1}{2}} V Z^{1/2})^{-1}$

$$= I - \frac{tZ^{-1/2} V Z^{1/2}}{1 + tr(tZ^{-1/2} V Z^{1/2})}$$

whereas $(Z + tV)^{-1} = Z^{-1} - \frac{tZ^{-1} V Z^{-1}}{1 + tr(tV)}$

$$= Z^{-1}\left(I - \frac{tZ^{-1/2} V Z^{-1/2}}{1 + tr(tV)}\right)$$

$$= Z^{-1}(I + tZ^{-1/2} V Z^{1/2})^{-1}$$

Thus, $g(t) = tr\left[Z^{-1}(I + tZ^{-\frac{1}{2}} V Z^{1/2})\right]$

Let $Z^{1/2} V Z^{1/2} = Q \Lambda Q^T$. Then,

$$g(t) = tr\left[Z^{-1}(I + tQ\Lambda Q^T)^{-1}\right]$$

$$tr(Z^{-1}(I + tQ\Lambda Q^T)^{-1}) = tr\left(Z^{-1}\left(I - \frac{tQ\Lambda Q^T}{1 + tr(tQ\Lambda Q^T)}\right)\right)$$

$$= tr(Z^{-1}(Q(I + t\Lambda)^{-1} Q^T))$$

$$= tr\left(Z^{-1}\left(QQ^T - \frac{t\Lambda}{1 + tr(t\Lambda)}\right)\right)$$

Thus, $g(t) = tr\left(Z^{-1} Q (I + t\Lambda)^{-1} Q^T\right)$

$= tr\left(Q^T Z^{-1} Q (I + t\Lambda)^{-1}\right)$

$= \sum_i \left[(Q^T Z^{-1} Q)_{ii} \, (1 + t\lambda_i)^{-1} \right]$

But $Z \succ 0 \Rightarrow (Q^T Z^{-1} Q)_{ii} > 0 \;\; \forall i$

and $(1 + t\lambda_i)^{-1}$ is convex for $\forall i$

Thus $g(t)$ is convex for $\forall t \in \mathbb{R}_+$

$\Rightarrow f(x)$ is convex.

$\boxed{\text{I } 3.19 \text{ (a)}}$

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.

$$\sum_{i=1}^{k} x_{[i]} \text{ is convex on } \mathbb{R}^n.$$

for $\forall \; 1 \leq k \leq r$.

Thus, $(\alpha_1 - \alpha_2) x_{[1]}$ is convex and $x_{[1]} + x_{[2]}$ is convex, which implies that $\alpha_2 (x_{[1]} + x_{[2]})$ is convex and $(\alpha_1 - \alpha_2) x_{[1]} + \alpha_2 (x_{[1]} + x_{[2]}) = \alpha_1 x_{[1]} + \alpha_2 x_{[2]}$ is convex. We can continue this process, inductively, as follows:

$\alpha_r \sum_{i=1}^{r} \alpha_i x_{[i]}$ is convex,

$(\alpha_{r-1} - \alpha_r) \sum_{i=1}^{r-1} \alpha_i x_{[i]}$ is convex

$$\Rightarrow \alpha_r \sum_{i=1}^{r} \alpha_i x_{[i]} + (\alpha_{r-1} - \alpha_r) \sum_{i=1}^{r-1} \alpha_i x_{[i]}$$

$$= \alpha_{r-1} \sum_{i=1}^{r-1} \alpha_i x_{[i]} + \alpha_r x_{[r]} \text{ is convex.}$$

This is the base case. For the inductive step, suppose that

$$\alpha_{j-1} \sum_{i=1}^{j-1} \alpha_i x_{[i]} + \sum_{i=j}^{r} \alpha_i x_{[i]} \text{ is convex}$$

for some j s.t. $1 \leq j \leq r$.

Then, $(\alpha_{j-2} - \alpha_{j-1}) \sum_{i=1}^{j-2} \alpha_i X_{[i]}$

$\alpha_{j-1} \sum_{i=1}^{j-1} \alpha_i X_{[i]} + \sum_{i=j}^{r} \alpha_i X_{[i]}$

$= \alpha_{j-2} \sum_{i=1}^{j-2} \alpha_i X_{[i]} + \sum_{i=j-1}^{r} \alpha_i X_{[i]}$ is convex

Thus, by (inverse) induction on j,

$\sum_{i=1}^{r} a_i X_{[i]}$ is convex.

$\boxed{A\ 2.10}$

$f(x) = (\Pi_n x_k)^{1/n}$, dom $f = \mathbb{R}^n_{++}$ is concave.
(proof on pg 74).

Let $f(x) = \prod_{k=1}^{n} x_k^{a_k}$, dom $f = \mathbb{R}^n_{++}$,

where $a_k \geq 0$, with $\sum_k a_k \leq 1$.

$$\frac{\partial f}{\partial x_k} = a_k x_k^{-1} f(x) \quad , \quad \frac{\partial^2 f}{\partial x_k^2} = a_k(a_k-1)\, x_k^{-2} f(x)$$

$$\frac{\partial^2 f}{\partial x_k \partial x_j} = a_k a_j x_k^{-1} x_j^{-1} f(x)$$

$$\Rightarrow \nabla^2 f(x) = A f(x), \text{ where } A_{ij} = \begin{cases} a_i(a_i-1)\, x_i^{-2} & \text{if } i=j \\ a_i a_j x_i^{-1} x_j^{-1} & \text{if } i \neq j \end{cases}$$

Let $y = 1/x$. we must show that
$v^T A v \leq 0$, $\forall\, v$.

$$v^T A v = \sum_{i,j} a_i a_j y_i y_j v_i v_j - \sum_i a_i y_i^2 v_i^2$$

$$= \underbrace{\tfrac{1}{2}\Big[\sum_i (a_i y_i v_i)\Big]^2 - \tfrac{1}{2}\sum (a_i y_i v_i)^2}_{\leq 0} - \underbrace{\sum_i a_i y_i^2 v_i^2}_{\leq 0}$$

The sum of the first two terms is less than zero by the Cauchy-Schwartz inequality. $\Rightarrow$ $v^T A v \leq 0$ $\Rightarrow$ $A \preceq 0$

$$\Rightarrow \nabla^2 f(x) = A f(x) = A \prod_{k=1}^{n} x_k^{a_k} \preceq 0.$$

$\Rightarrow$ $f(x)$ is concave.

$\boxed{\text{A 2.51 (a)}}$

$f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = -(\sqrt{x_1} + \cdots + \sqrt{x_n})^2$,
$\operatorname{dom} f = \mathbb{R}^n_{++}$.

$$-f(\theta x + (1-\theta)x) = \left(\sqrt{\theta x_1 + (1-\theta)x_1} + \cdots + \sqrt{\theta x_n + (1-\theta)x_n}\right)^2$$

$$\leq \left(|\theta x_1 + (1-\theta)x_1| + \cdots + |\theta x_n + (1-\theta)x_n|\right)$$

$$= \left[\left(|\theta x_1| + \cdots + |\theta x_n|\right) + \left(|(1-\theta)x_1| + \cdots + |(1-\theta)x_n|\right)\right]$$

$$= \left(|\theta x_1 + (1-\theta)x_1| + \cdots + |\theta x_n + (1-\theta)x_n|\right)$$

$= (|x_1| + \cdots + |x_n|)$. On the other hand,

$$-(f(\theta x) + f((1-\theta)x))$$

$$= \left(\sqrt{\theta x_1} + \cdots \sqrt{\theta x_n}\right)^2 + \left(\sqrt{(1-\theta)x_1} + \cdots + \sqrt{(1-\theta)x_n}\right)^2$$

$$= \theta\left(\sqrt{x_1} + \cdots + \sqrt{x_n}\right)^2 + (1-\theta)\left(\sqrt{x_1} + \cdots + \sqrt{x_n}\right)^2$$

$$= \left(\sqrt{x_1} + \cdots + \sqrt{x_n}\right)^2 = f(x).$$

$$\frac{\partial f}{\partial x_1} = 2 \cdot \frac{1}{2} x_i^{-\frac{1}{2}} \left(\sqrt{x_1} + \cdots + \sqrt{x_n}\right)$$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{1}{2} x_i^{-\frac{1}{2}} x_j^{-\frac{1}{2}} \left(\sqrt{x_1} + \cdots + \sqrt{x_n}\right)$$

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{1}{4} x_i^{-\frac{3}{2}} \left(\sqrt{x_1} + \cdots + \sqrt{x_n}\right) \qquad ?$$

No luck...

contd. on next page...

A diffrent approach:

$\boxed{A\ 2.51\ (a)}$

$f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = -(\sqrt{x_1} + \cdots + \sqrt{x_n})^2$,
$\operatorname{dom} f = \mathbb{R}^n_{++}$.

$f(x) = -(y_1 + \cdots + y_n)^2$, where $y_i = \sqrt{x_i}$.

Thus, $f(x) = h(g(x))$, where $g(x) = \sqrt{x}$
and $h(x) = -(x_1 + \cdots + x_n)^2$.

h is concave and non-increasing and
g is convex. Thus, f is concave.

$\boxed{A\ 5.8}$

$$\min \sum_{k=1}^{N} (x^T g(t_k) - y_k)^2$$

s.t. $x^T g(t)$ is convex in t on $[\alpha_0, \alpha_M]$.

(a.) Express in the form
$\min \|Ax-b\|_2^2$
s.t. $Gx \le h$.

We need $\frac{\partial^2}{\partial t^2} x^T g(t) \ge 0$.

Let $g(t) = at^3 + bt^2 + ct + d$, $a, b, c, d \in \mathbb{R}^m$.

$\frac{\partial^2}{\partial t^2} g(t) = 6at + 2b$

$x^T(3at + b) \ge 0$

$\Leftrightarrow -(3ta + b)^T x \le 0$

That is, $Gx \le h$, with $G = -(3ta+b)^T$, $h = 0$

We have:
$$\boxed{\begin{aligned} &\min \|g^T(t)x - y\|_2^2 \\ &\text{s.t. } -(3ta+b)^T x \le 0 \end{aligned}}$$

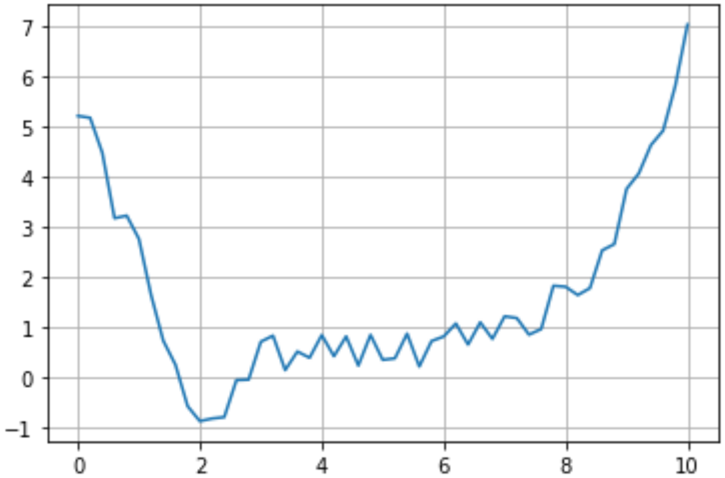
$(G \in \mathbb{R}^{11 \times 13})$

$G = -\begin{bmatrix} - g_{pp,0} - \\ - g_{pp,1} - \\ \vdots \\ - g_{pp,M} - \end{bmatrix}$ $\quad - g_{pp}^T x \le 0$.

$A = \begin{bmatrix} - g_0 - \\ \vdots \\ - g_M - \end{bmatrix}$

In [11]:
```
import numpy as np
import cvxpy as cp
import matplotlib.pyplot as plt
```

In [13]:
```
from spline_data import t, y
plt.plot(t,y)
plt.grid()
plt.show()
```



In [44]:
```
from bsplines import bsplines

G = np.zeros((11,13))

i = 0
for u in np.arange(11):
    _, _, gpp = bsplines(u)
    G[i,:] = -gpp
    i += 1
```

In [45]:
```
A = np.zeros((51,13))

i = 0
for ti in t:
    g,_,_ = bsplines(ti)
    A[i,:] = g
    i += 1
```

In [46]:
```
# Define and solve the CVXPY problem.
M = 13
x = cp.Variable(M)
cost = cp.sum_squares(A @ x - y)
prob = cp.Problem(cp.Minimize(cost),[G @ x <= 0])

print("\nThe optimal value is", prob.solve())
print("The solution x is")
print(x.value)
```

```

The optimal value is 7.844981009336883
The solution x is
[ 5.69438888  4.41060672  1.8430424  -0.38409468 -0.08762342  0.20884784
  0.5053191   0.80179036  1.09826161  1.39473287  3.12555339  5.22124366
  6.90143774]
```

In [53]:
```
g_s = np.matmul(A,x.value)
plt.plot(t,g_s)
plt.plot(t,y)
plt.grid()
plt.show()
```