EECS 551 Homework 7 YUZHAN JIANG

P1.

(a) $\|A\|_2 = \max_{x\neq 0} \frac{\|Ax\|_1}{\|x\|} = \sigma_1$

$\|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2}$

So that $\sigma_1^2 \leq \sigma_1^2 + \sigma_2^2 + \dots + \sigma_r^2$

$\Rightarrow \sqrt{\sigma_1^2} \leq \sqrt{\sigma_1^2 + \sigma_2^2 + \dots \sigma_r^2}$

$\sigma_1 \leq \sqrt{\sum_{i=1}^{r} \sigma_i^2}$

$\therefore \|A\|_2 \leq \|A\|_F$

(b) $\|A\|_F = \sqrt{\sum_{i}^{r} \sigma_i^2} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_r^2} \leq \sqrt{\sigma_1^2 + \sigma_1^2 + \dots + \sigma_1^2}$ (Since $\sigma_1 > \sigma_2 > \dots > \sigma_r$)

$\leq \sqrt{r \cdot \sigma_1^2}$

$\leq \sqrt{r}\, \sigma_1$

$\leq \sqrt{r}\, \|A\|_2$

$\therefore \|A\|_F \leq \sqrt{r}\, \|A\|_2$

(c) $\|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_r^2}$

$\leq \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_r^2 + 2\sigma_1\sigma_2 + 2\sigma_1\sigma_3 + \dots + \dots 2\sigma_{r-1}\sigma_r}$

$= \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + 2\sum_{i\neq j} \sigma_i \sigma_j}$

$= \sqrt{(\sigma_1 + \sigma_2 + \dots + \sigma_r)^2}$

$= \sigma_1 + \sigma_2 + \dots + \sigma_r$

$= \|A\|_*$

$\therefore \|A\|_F \leq \|A\|_*$

f) Let $y \in F^m$ $\quad \|y\|_2^2 = \sum_{i=1}^{m} |y_i|^2 \leq m \cdot (\max_j |y_j|^2) = m \cdot \|y\|_\infty$

$\Rightarrow \|y\|_2^2 \leq m \|y\|_\infty$

$\|A\|_2 = \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \leq \max_{x\neq 0} \frac{\sqrt{m}\, \|Ax\|_\infty}{\|x\|_2}$

$= \sqrt{m} \max \frac{\|Ax\|_\infty}{\|x\|_2}$

$= \sqrt{m}\, \|A\|_\infty$

Therefore, $\|A\|_2 \leq \sqrt{m}\, \|A\|_\infty$

P2:

First construct compact SVD of $A = U_r \Sigma_r V_r'$ and $B = X_s \Omega_s Y_s'$ and $r = \text{rank}(A)$ and $s = \text{rank}(B)$

If $AB' = 0$, $\Rightarrow U_r \Sigma_r V_r' Y_s \Omega_s' X_s' = 0$

$$\Rightarrow (\Sigma_r^{-1} U_r') U_r \Sigma_r V_r' Y_s \Omega_s' X_s' (X_s \Omega_s^{-1}) = 0$$

$$\Rightarrow V_r' Y_s = 0$$

Similarly if $A'B = 0$ $\Rightarrow$ $(U_r \Sigma_r V_r')' X_s \Omega_s Y_s' = 0$

$$V_r \Sigma_r' U_r' X_s \Omega_s Y_s' = 0$$

$$\Rightarrow (\Sigma_r')^{-1} V_r' V_r U_r' X_s \Omega_s Y_s' (Y_s \Omega_s^{-1}) = 0$$

$$\Rightarrow U_r' X_s = 0$$

Thus, $A+B = U_r \Sigma_r V_r' + X_s \Omega_s Y_s'$

$$= [U_r \quad X_s] \begin{bmatrix} \Sigma_r & 0 \\ 0 & \Omega_s \end{bmatrix} [V_r \quad Y_s]'$$

$$\therefore \|A+B\|_* = \|\Sigma_r\|_* + \|\Omega_s\|_*$$

$$= \|A\|_* + \|B\|_*$$

P3:

(a) let $A = U_r \Sigma_r V_r'$, $A^+ = V_r \Sigma_r^{-1} U_r'$

and let P be a permutation matrix

$$\Rightarrow A^+ = (V_r \cdot P')(P \cdot \Sigma_r^{-1} P')(P U_r')$$

$$\|A^+\|_2 = \sigma_1(A^+) = \|P \Sigma_r^{-1} P'\|_\infty$$

$$= \|\Sigma_r^{-1}\|_\infty$$

$$= \frac{1}{\sigma_r}$$

$\therefore$ Generalized inverse matrix G is with spectral norm is $\frac{1}{\sigma_r}$

(b) let $A = U \Sigma V'$

The generalized inverse matrix $G = V \begin{bmatrix} \Sigma_r^{-1} & S_2 \\ S_3 & S_4 \end{bmatrix} U'$ where $\Sigma_r^{-1} = \begin{bmatrix} \frac{1}{\sigma_r} & & \\ & \ddots & \\ & & \frac{1}{\sigma_1} \end{bmatrix}$

S_4 is the diagonal matrix with non-zero elements are 7

$\therefore$ G has spectral norm 7

P4.

proof:

We are given that A and B are orthogonal projection matrix,

$AA = A$ and $A' = A = AA$

$BB = B$, and $B' = B = BB$

and the eigenvals of A and B are either 0 or 1

$$\left(\tfrac{1}{2}A+\tfrac{1}{2}B\right)'\left(\tfrac{1}{2}A+\tfrac{1}{2}B\right) = \tfrac{1}{4}A'A + \tfrac{1}{4}B'A + \tfrac{1}{4}A'B + \tfrac{1}{4}B'B$$
$$= \tfrac{1}{2}AA' + \tfrac{1}{4}BA' + \tfrac{1}{4}AB' + \tfrac{1}{4}BB'$$
$$= \left(\tfrac{1}{2}A+\tfrac{1}{2}B\right)\left(\tfrac{1}{2}A+\tfrac{1}{2}B\right)'$$

$\therefore$ $P\left(\tfrac{1}{2}A+\tfrac{1}{2}B\right) = \left\|\tfrac{1}{2}A+\tfrac{1}{2}B\right\|_2 \leq \tfrac{1}{2}\|A\|_2 + \tfrac{1}{2}\|B\|_2 = \tfrac{1}{2}+\tfrac{1}{2} = 1$

$\therefore$ the spectral radius of their average, $\tfrac{1}{2}A+\tfrac{1}{2}B$ is at most 1.

P5:

(a) $\|f(x) - f(y)\|_\beta = \| A^T(Ax-b) - A^T(Ay-b)\|_\beta$

$$= \| A^T(Ax-b-Ay+b)\|_\beta$$

$$= \| A^T(Ax-Ay)\|_\beta$$

$$= \| A^T A(x-y)\|_\beta$$

$$\leq L\|x-y\|_\alpha$$

$$L \geq \frac{\|A^T A(x-y)\|_2}{\|x-y\|_2}$$

$$\Rightarrow \quad L \geq \max \frac{\|A^T A(x-y)\|_2}{\|x-y\|_2}$$

$$= \|A^T A\|_2$$

$$= \sigma_1^2(A)$$

$\therefore$ The best Lipschitz constant is σ_1^2

(b) We are given that $\alpha = \beta = 2$

Let $f(x) = \|x\|_2$ and $f(y) = \|y\|_2$

So that $f(x) = \|x\|_2 = \|x - Y + Y\|_2$

$$\leq \|x-Y\|_2 + \|Y\|_2$$

$$= \|x-Y\|_2 + f(y)$$

$$\therefore \quad f(x) - f(y) \leq \|x-Y\|_2$$

$$\|f(x) - f(y)\|_2 \leq \|x-Y\|_2$$

Where here $L = 1$

P6:

(b)

In [14]:
```julia
#Question b
using Random: seed!
using Statistics: mean
seed!(0)
m = 100; n = 50; sigma = 2.0
A = randn(m,n)
xtrue = rand(n) # note that xtrue is non-negative
b = A * xtrue + sigma * randn(m)
x0 = A \ b; @show count(x0 .< 0), minimum(x0) # negative values
x0[x0 .<= 0] .= mean(x0[x0 .> 0]) # reasonable initial nonnegative guess

s = svdvals(A)
mu = 1/(s[1,1]^2)
x_100 = nnlsgd(A, b ; mu, x0=x0,  nIters=400)
x_100[1:3]
```

```
(count(x0 .< 0), minimum(x0)) = (3, -0.45686878247066276)
```

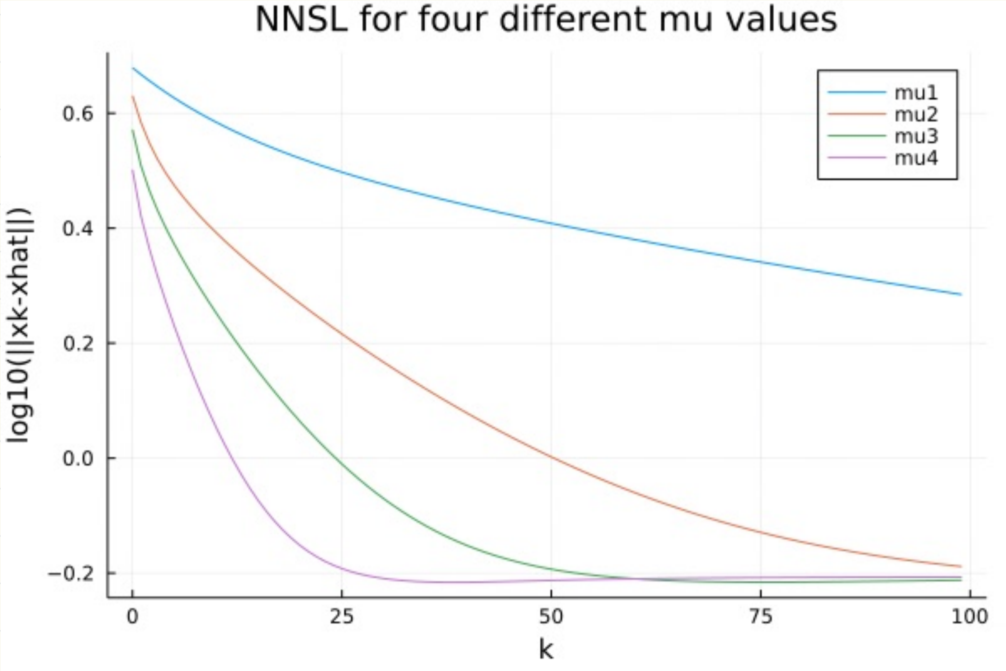
Out[14]:
```
3-element Vector{Float64}:
 0.890131225634076
 0.9796146939170665
 1.2012483516740196
```

(c)

In [15]:
```julia
#Question c
using Optim # you will likely need to add this package
using LinearAlgebra
lower = zeros(n)
upper = fill(Inf, (n,))
inner_optimizer = GradientDescent()
f = x -> 1/2 * norm(A * x - b)^2 # cost function
function grad!(g, x) # its gradient
g[:] = A' * (A * x - b)
end
results = optimize(f, grad!, lower, upper, x0,Fminbox(inner_optimizer), Optim.Options(g_tol=1e-12))
xnnls = results.minimizer
xnnls[5:7]
```

Out[15]:
```
3-element Vector{Float64}:
 1.058567280194555
 0.372060621615 7402
 0.9069908026761518
```

(d)



NNSL for four different mu values
mu1
mu2
mu3
mu4
log10(||xk-xhat||)
0.6
0.4
0.2
0.0
−0.2
0
25
50
75
100
k

P7.

(a) $$a_*, \mu_*, Q_* = \underset{a\in R, \mu\in R^d, Q: Q'Q=I_d}{\arg\min} \| B - a Q (A - \mu_A 1_n') - \mu 1_n' \|_F$$

$$\mu_* = (1_n)^+ (B - aQ(A - \mu_A 1_n'))$$
$$= \frac{1}{n} 1_n B - \frac{a}{n} \cdot (1_n A - 1_n \mu_A 1_n') Q$$
$$= \mu_B - \frac{a}{n} (n\mu_A - n\mu_A) Q$$
$$= \mu_B$$

and $$Q_* = \underset{Q: Q'Q=I_d}{\arg\min} \| B_0' - Q'(aA_0)' \|_F$$ with $\mu_* = \mu_B$ and fixed value a

We know that $U\Sigma V'$ is the SVD of $B_0 A_0'$

$$Q^* = UV'$$

Finally when $\mu_* = \mu_B$ and $Q_* = UV'$ and fixed a

$$a_* = \min_a \| B_0 - A_0 Q^* a \|_F$$

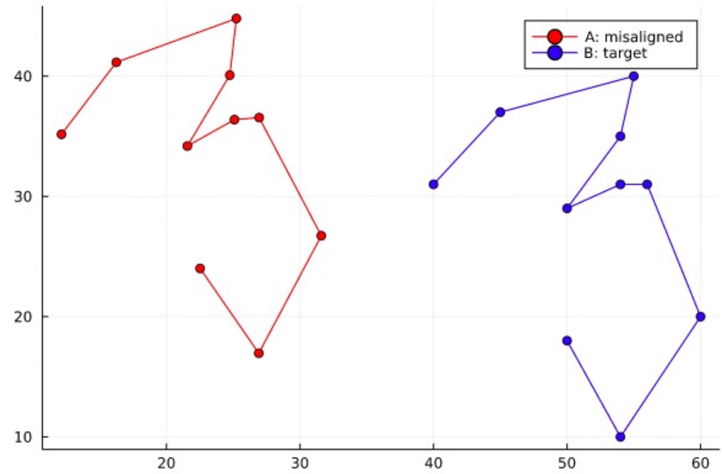
$$= Tr((B_0 - A_0 Q_* a)'(B_0 - A_0 Q_*))$$
$$= a^2 \, Tr((A_0 Q_*)'(A_0 Q_*)) - 2a Tr(B_0 A_0 Q_*) + Tr(B_0' B_0)$$

Let derivative $= 0$,

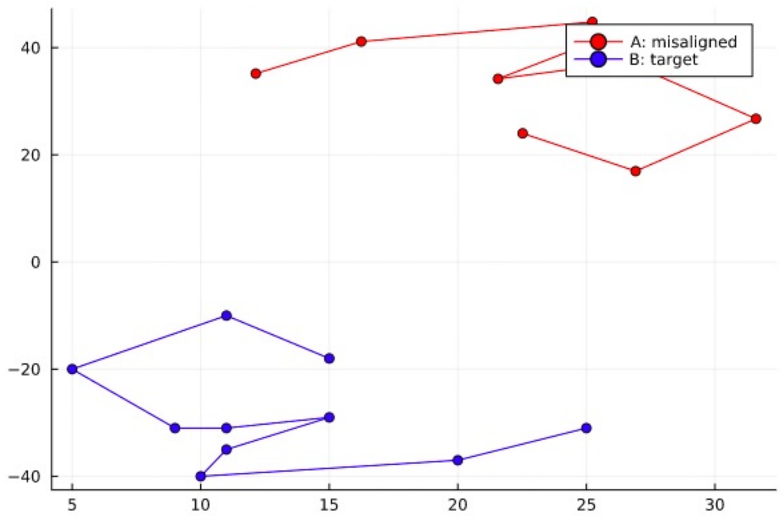
$$\Rightarrow 2a \, Tr((A_0 Q_*)'(A_0 Q_*)) - 2Tr(B_0 (A_0 Q_*)') = 0$$
$$\Rightarrow 2a \, Tr(Q_*' A_0' A_0 Q_*) = 2Tr(B_0 A_0' Q_*')$$
$$2a \, Tr(A_0' A_0 Q_* Q_*') = 2Tr(B_0 A_0' Q_*')$$
$$2a \, Tr(A_0' A_0) = 2Tr(B_0 A_0' Q_*')$$
$$\therefore \quad a_* = \frac{Tr(B_0 A_0' Q_*')}{Tr(A_0 A_0')}$$

(c)

Out[19]:



Out[17]:



(d) The resulting error $\|\hat{A}-B\|_F$ for data 1 is 91.807

The resulting error $\|\hat{A}-B\|_F$ for data 2 is 208.001