MSSC 5931 - Homework 1

Henri Medeiros Dos Reis

October 15, 2022

1. **Iteratively reweighted least squares for 1-norm approximation.**

   Solution 1. .

   (a)

$$x_{ls} = (A^TA)^{-1}Ay$$

$$x_{wls} = \sum_{i=1}^{m} w_i(a_i^Tx - y_i)^2 = \sum_{i=1}^{m}(\sqrt{w_i}a_i^Tx - \sqrt{w_i}y_i)^2 \Rightarrow$$

$$\left|\left|\begin{bmatrix} \sqrt{w_1}a_1^Tx - \sqrt{w_1}y_1 \\ \sqrt{w_2}a_2^Tx - \sqrt{w_2}y_2 \\ \vdots \\ \sqrt{w_m}a_m^Tx - \sqrt{w_m}y_m \end{bmatrix}\right|\right|^2 = \left|\left|\begin{bmatrix} \sqrt{w_1}a_1^T \\ \sqrt{w_2}a_2^T \\ \vdots \\ \sqrt{w_m}a_m^T \end{bmatrix} x - \begin{bmatrix} \sqrt{w_1}y_1 \\ \sqrt{w_2}y_2 \\ \vdots \\ \sqrt{w_m}y_m \end{bmatrix}\right|\right|^2$$

Let W and W be diagonal matrices such that

$$W = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_m \end{bmatrix}, \tilde{W} = \begin{bmatrix} \sqrt{w_1} & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{w_m} \end{bmatrix}$$

Then $||\tilde{W}Ax - \tilde{W}y||^2$ and let $\tilde{A}$ be $\tilde{W}A$, hence

$$x_{wls} = ||\tilde{A} - \tilde{y}||^2 = (\tilde{A}^T\tilde{A})^{-1}\tilde{A}^T\tilde{y} = [(\tilde{W}A)^T(\tilde{W}A)]^{-1}(\tilde{W}A)^T\tilde{W}y \Rightarrow$$

$$[A^TW^T\tilde{W}A]^{-1}A^T\tilde{W}^T\tilde{W}y, \text{ Since } \tilde{W}^T\tilde{W} = W \text{ then}$$

$$x_{wls} = (A^TWA)^{-1}A^TWy$$

   (b) For the cost function in (3) to be equal to the l_1-norm approximation error, then we need to set them equal to each other and solve for the weights.

$$\sum_{i=1}^{m} w_i(x)(\tilde{a}_i^Tx - y_i)^2 = \sum_{i=1}^{m} |\tilde{a}_i^Tx - y_i|$$

Then solve for $w_i(x)$, looking at each iteration

$$w_i(x) = \left|\frac{a_i^T x - y_i}{(a_i^T x - y_i)^2}\right| = \left|\frac{1}{a_i^T x - y_i}\right|$$

Therefore, to minimize the cost function we use the weights by using $w_i(x) = \left|\frac{1}{a_i^T x - y_i}\right|$

2. **Estimating a signal with interference.**

   Solution 2.

3. **Identifying a system from input/output data.**

   Solution 3. .

   (a) Let a_i be the i^{th} row of A, so that we can express this as a double sum.

$$J = \sum_{k=1}^{N}\sum_{i=1}^{m}\left(a_i^T x^{(k)} - y^{(k)}\right)^2$$

Then we need to minimize the inner sum, which is only dependent on a_i. Let's look at only that sum. And call that J_i

$$J_i = \sum_{i=1}^{m}\left(a_i^T x^{(k)} - y^{(k)}\right) = \sum_{i=1}^{m}\left(x_i^{(k)T} a_i - y^{(k)}\right) \Rightarrow$$

$$J_i = \left|\left|\begin{bmatrix} x^{(1)T} \\ x^{(2)T} \\ \vdots \\ x^{(k)T} \end{bmatrix} a_i - \begin{bmatrix} y_i^{(1)T} \\ y_i^{(2)T} \\ \vdots \\ y_i^{(k)T} \end{bmatrix}\right|\right| = \left|\left|X^T a_i - y_i\right|\right|$$

Which has the same form of least squares estimation. Where we want to chose an estimate of $\hat{a}_i$ that minimizes J_i, therefore

$$\hat{a}_i = (X^T X)^{-1} X^T y_i, \text{ where } y_i = \begin{bmatrix} y_i^{(1)T} \\ y_i^{(2)T} \\ \vdots \\ y_i^{(k)T} \end{bmatrix}$$

Now using $\hat{a}_i$ and Y, we can estimate $\hat{A}$. Since $Y = \begin{bmatrix} Y^{(1)} & Y^{(2)} & \cdots & Y^{(N)} \end{bmatrix}$, and we are using y_i as a column vector, plugging back to our previous equation will give us $\hat{A}^T$

$$\hat{A}^T = (X^T X)^{-1} X Y^T, \text{ then we take the transpose:}$$

$$\hat{A} = ((X^T X)^{-1} X Y^T)^T = (Y^T)^T X^T ((X^T X)^{-1})^T = Y X^T (X X^T)^{-1}$$

(b)

```
# Compute the inverse of XX^T and the A_hat matrix
invXXT = solve(X%*%t(X))
A_hat  = Y%*%t(X)%*%invXXT

N <- ncol(X)

# Set up the needed variables
norm(A_hat%*%X[,1]-Y[,1], 2) #top
norm(Y[,1], 2) #bottom
sum<-0

# Take the sum and divide by N
for (i in c(1:N))
{
  top<-norm(A_hat%*%X[,i]-Y[,i], 2) #top
  bot<-norm(Y[,i], 2) #bottom
  sum<-sum+ (top/bot)
}
(result <- sum/N)

#---------Output-----------
#0.05814324
```

4. **Robust input design.**

Solution 4. .

a-) Since x^{ln} is going too have form of least squares. We can use the average, because it is a good estimation, and $y = Ax$, then we can switch the variables such that $y^{des} = \bar{A}x$. Where $\bar{A} = \frac{1}{K}\sum_{i=1}^{K} A^{(i)}$.

$$x^{ln} = \bar{A}^T(\bar{A}\bar{A}^T)^{-1}y^{des} = \left(\frac{1}{K}\sum_{i=1}^{K} A^{(i)^T}\right)\left[\frac{1}{K}\sum_{i=1}^{K} A^{(i)}\frac{1}{K}\sum_{i=1}^{K} A^{(i)^T}\right]^{-1} y^{des}$$

And for the mean square error minimization

$$\frac{1}{K}\sum_{i=1}^{K} ||y^{(i)} - y^{des}||^2 = \frac{1}{K}\sum_{i=1}^{K} ||A^{(i)}x - y^{des}||^2$$

Then, when we expand the sum

$$\frac{1}{K}(||A^{(1)}x - y^{des}||^2 + ||A^{(2)}x - y^{des}||^2)\cdots||A^{(K)}x - y^{des}||^2$$

Which has the same form as minimizing weighted sum objective. Therefore we can write this as

$$\left\| \frac{1}{\sqrt{K}} \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(K)} \end{bmatrix} x - \frac{1}{\sqrt{K}} \begin{bmatrix} y^{des} \\ y^{des} \\ \vdots \\ y^{des} \end{bmatrix} \right\| = \left\| \frac{1}{\sqrt{K}} \tilde{A} x - \frac{1}{\sqrt{K}} \tilde{y}^{des} \right\|$$

Where $\tilde{A} = \frac{1}{\sqrt{K}} \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(K)} \end{bmatrix}$ and $y^{des} = \begin{bmatrix} y^{des} \\ y^{des} \\ \vdots \\ y^{des} \end{bmatrix}$ and let $\tilde{A}$ be full rank.

$$x^{mmse} = \left(\frac{1}{\sqrt{K}} \tilde{A}^T \frac{1}{\sqrt{K}} \tilde{A} \right)^{-1} \left(\frac{1}{\sqrt{K}} \tilde{A} \frac{1}{\sqrt{K}} \tilde{y}^{des} \right) = \left(\frac{1}{K} \tilde{A}^T \tilde{A} \right)^{-1} \left(\frac{1}{K} \tilde{A} \tilde{y}^{des} \right) \Rightarrow$$

$$\left(\tilde{A}^T \tilde{A} \right)^{-1} \left(\tilde{A} \tilde{y}^{des} \right) \Rightarrow$$

$$\left[(A^{(1)T} A^{(1)} + A^{(2)T} A^{(2)} + \cdots + A^{(K)T} A^{(K)})^{-1} (A^{(1)} y^{des} + A^{(2)} y^{des} + \cdots + A^{(K)} y^{des}) \right] \Rightarrow$$

$$x^{mmse} = \left(\sum_{i=1}^{K} (A^{(i)T} A^{(i)})^{-1} \right) \left(\sum_{i=1}^{K} (A^{(i)} y^{des}) \right)$$

b-)

```
A_bar <- 1/K*rowSums(A, dim=2)
x_ln <-  t(A_bar)%*%solve((A_bar%*%t(A_bar)))%*%y_des


total_left <- 0
total_right <- 0
for (i in 1:K)
{
  total_left <- total_left+ (t(A[,,i])%*%A[,,i])
  total_right <- total_right+(t(A[,,i])%*%y_des)

}
total_left <- solve(total_left)
x_mmse <- total_left%*%total_right

#residual least norm
norm(A_bar%*%x_ln -y_des, 2)

#residual mean squared
```

```
y_tilde_des <- rep(y_des, K)
A_tilde<-A[,,i]
for(i in 2:K)
{
  A_tilde <- rbind(A_tilde, A[,,i])
}
1/sqrt(K)*norm(y_tilde_des - A_tilde%*%x_mmse, 2)
```

5. **Householder reflections.**

   Solution 5. .

a-) For Q to be orthogonal $Q^TQ = I$

$$Q^TQ = (I - 2uu^T)^T(I - 2uu^T) \Rightarrow$$

$$(I - 2(u^T)^Tu^T)(I - 2uu^T) = (I - 2uu^T)(I - 2uu^T) \Rightarrow$$

$$I - 4uu^T + 4(uu^T)(uu^T) = I - 4uu^T + 4u(u^Tu)u^T, \text{ and since } u^Tu = 1$$

$$I - 4uu^T + 4uu^T = I$$

Therefore Q is orthogonal.

b-) First, lets show $Qu = -u$:

$$-u = (I - 2uu^T)u = u - 2uu^Tu = u - 2u(u^Tu) = u - 2u = -u \text{ since} u^Tu = 1$$

Second, $Qv = v$ where $u^Tv = 0$:

$$Qv = (I - 2uu^T)v = v - 2u(u^Tv) = v - 2u(0) = v$$

c-) Let's start by multiplying Qx, then substituting u with $\frac{v}{||v||}$ and v with $(x + \alpha e_1)$

$$Qx = (I - 2uu^T)x = x - 2uu^Tx = x - 2\left(\frac{v}{||v||}\right)\left(\frac{v}{||v||}\right)^T x \Rightarrow$$

$$x - 2\left(\frac{x + \alpha e_1}{||x + \alpha e_1||}\right)\left(\frac{x + \alpha e_1}{||x + \alpha e_1||}\right)^T x = x - 2\left(\frac{x + \alpha e_1}{||x + \alpha e_1||}\right)\left(\frac{x^Tx + \alpha e_1^Tx}{||x + \alpha e_1||}\right) \Rightarrow$$

$$x - 2\left(\frac{x + \alpha e_1}{||x + \alpha e_1||}\right)\left(\frac{||x||^2 + \alpha x_1}{||x + \alpha e_1||}\right) = x - 2\frac{x||x||^2 + \alpha e_1||x||^2 + \alpha x_1 x + \alpha^2 e_1 x_1}{||x + \alpha e_1||} \Rightarrow$$

$$x - 2\frac{x||x||^2 + \alpha e_1||x||^2 + \alpha x_1 x + \alpha^2 e_1 x_1}{(x + \alpha e_1)^T(x + \alpha e_1)} = x - 2\frac{x||x||^2 + \alpha e_1||x||^2 + \alpha x_1 x + \alpha^2 e_1 x_1}{||x||^2 + x^T\alpha e_1 + \alpha e_1^Tx + \alpha e_1^T\alpha e_1} \Rightarrow$$

$$\frac{x||x||^2 + xx_1\alpha + x\alpha x_1 + x\alpha e_1^T\alpha e_1}{||x||^2 + x^T\alpha e_1 + \alpha e_1^Tx + \alpha e_1^T\alpha e_1} - 2\frac{x||x||^2 + \alpha e_1||x||^2 + \alpha x_1 x + \alpha^2 e_1 x_1}{||x||^2 + x^T\alpha e_1 + \alpha e_1^Tx + \alpha e_1^T\alpha e_1} \Rightarrow$$

$$\frac{-x||x||^2 + x\alpha x_1 + x\alpha x_1 + x\alpha^2 - 2x||x||^2 - 2\alpha x_1 x - 2\alpha^2 e_1 x_1}{||x||^2 + 2\alpha x_1 + \alpha^2} \Rightarrow$$

$$\frac{-x||x||^2 + x\alpha^2 - 2\alpha e_1||x||^2 - 2\alpha^2 e_1 x_1}{||x||^2 + 2\alpha x_1 + \alpha^2}$$

The only terms that do not have e_1 on the numerator are $-x||x||^2 + x\alpha^2$, so let's take those and set them equal to zero. So that we only have terms that depend on e_1. Which will make this whole equation lie on the line through e_1.

$$-x||x||^2 + x\alpha^2 = 0 \Rightarrow x\alpha^2 = x||x||^2 \Rightarrow \alpha^2 = ||x||^2 \Rightarrow \alpha = ||x||$$

Now substituting $alpha = ||x||$ in the equation:

$$\frac{-x||x||^2 + x||x||^2 - 2e_1||x||||x||^2 - 2e_1||x||^2 x_1}{||x||^2 + 2||x||x_1 + ||x||^2} \Rightarrow$$

$$\frac{-2e_1||x||||x||^2 - 2e_1||x||^2 x_1}{||x||^2 + 2||x||x_1 + ||x||^2} = \frac{-2||x||^2 e_1(||x|| + x_1)}{2||x||^2 + 2||x||x_1} \Rightarrow$$

$$\frac{-2||x||^2 e_1(||x|| + x_1)}{2||x||(||x|| + x_1)} = \frac{-e_1||x||^2}{||x||} = -||x||e_1$$

Now lets use the formulas we just solved in R so that we can compute the results for $x = (3, 2, 4, 1, 5)$. Then plug that Householder reflection to x to find Qx

```
X<-as.matrix(c(3,2,4,1,5),ncol=1)
alpha <- norm(X,2)
e_1<- as.matrix(c(1,0,0,0,0),ncol=1)
v <- X+alpha*e_1
u <- v/norm(v, 2)

#normal method
cat(Qx = ,Qx <- (diag(length(u)) - 2*(u%*%t(u)))%*%X)
#created method
cat(Qx_alt = ,Qx_alt <- -e_1%*%norm(X, 2))

#---------Output---------------
#Qx =  -7.416198 4.440892e-16 8.881784e-16 2.220446e-16
   4.440892e-16
#Qx_alt =  -7.416198 0 0 0 0
```

6. **True/false questions about linear algebra.**

Solution 6. .

a-) True

Since Q has orthonormal columns, we can separate it in 2 block matrices.
$R = [Q\ \tilde{Q}]$ Then

$$||R^T w||^2 = (R^T w)^T (R^T w) = w^T R R^T w = w^T w = ||w||^2$$

And we can also write this as

$$||R^T w||^2 = \left|\left| \begin{bmatrix} Q^T w \\ \tilde{Q}^T w \end{bmatrix} \right|\right|^2 = ||Q^T w||^2 + ||\tilde{Q}^T w||^2$$

Finally, by combining these two equations, we have

$$||w||^2 = ||Q^T w||^2 + ||\tilde{Q}^T w||^2 \Rightarrow ||w||^2 - ||\tilde{Q}^T w||^2 = ||Q^T w||^2$$

Therefore $||w|| \geq ||Q^T w||$.

b-) True

Since $dim(range(A)) + dim(null(A)) = p$, $null(A) = \{0\}$ and $dim(range(B)) + dim(null(B)) = q$, then

$$dim(range(A)) = p, \text{ and } dim(range(A)) \leq dim(range(B))$$

Therefore, $dim(range(A)) \leq dim(range(B)) + dim(null(B))$, which implies $p \leq q$.

c-) True

Since $V = [V_1\ V_2]$ and V is full rank, then V_1 and V_2 are linearly independent. Then $V_2 x \notin range(V_1)$. Then $AV_2 x = 0$ if either $V_2 x \in null(A)$ or $x \in null(A)$.
By using rank-nullity theorem.

$$dim(range(V_2)) + dim(null(V_2)) = n$$

Where $dim(null(V_2))$ has to be 0, therefore $dim(null(V_2)) = n$ Finally, if V is invertible, and $range(V_1) = null(A)$ then $null(AV_2) = \{0\}$.

d-) True

Since $rank(A) = rank(B) = rank([A\ B])$, then they have to lie in the same dimension of vectors spanned by their columns.
Which means that $range(A) \leq range(B)$, also, $range(B) \leq range(A)$.
Therefore $range(A) = range(B)$.

e-) False

Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $A^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ $null(A)$ is the y-axis, the range of A is the x-axis.
If we let $x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, then x is not in the range of A, and it is not in the null space of A^T.

f-) True

Since A is invertible, then

$$rank(B) = rank(AA^{-1}B) \leq rank(AB) \leq rank(B)$$

Therefore AB is not full rank if and only if B is not full rank.

g-) True

$dim(range(A)) + dim(null(A)) = n$, since we know that $dim(range(A)) \leq n$ because A is not full rank. Then $dim(null(A)) \geq 0$ which shows that there is a set of all n-dimensional vectors x such that $Ax = 0$

7. **Least-squares residuals.**

Solution 7. First, we need the residual vector to be perpendicular to Ax_{ls}.

$$Ax_{ls} = AA^Ty = A(A^TA)^{-1}A^Ty = y_{ls}$$

Then:

$$r = y - y_ls \Rightarrow$$

$$y_{ls}^Tr = y_{ls}^T(y - y_ls) = y_{ls}^Ty - y_{ls}^Ty_{ls} \Rightarrow$$

$$= y^T(A(A^TA)^{-1}A^T)^Ty - y^T(A(A^TA)^{-1}A^Ty)^T(A(A^TA)^{-1}A^T)y \Rightarrow$$

$$= y^T(A^T)^T(A^TA)^{-1}A^Ty - y^T(A^T)^T(A^TA)^{-1}A^TA(A^TA)^{-1}A^Ty \Rightarrow$$

$$= y^TA(A^TA)^{-1}A^Ty - y^TA[(A^TA)^{-1}A^TA](A^TA)^{-1}A^Ty$$

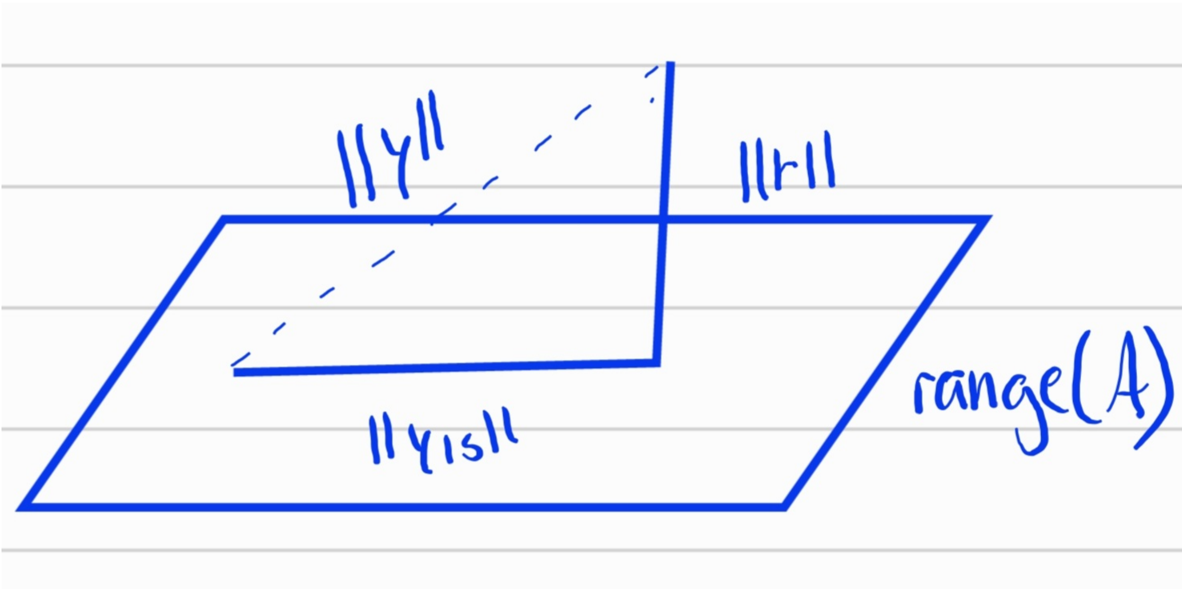
And since $(A^TA)^{-1}A^TA = I$, then

$$y^TA(A^TA)^{-1}A^Ty - y^TA(A^TA)^{-1}A^Ty = 0 = y_{ls}^Tr$$

Therefore they are perpendicular. Which finally leads us to show that

$$||y||^2 = ||y_{ls}||^2 + ||r||^2 \Rightarrow ||r||^2 = ||y||^2 - ||y_{ls}||^2$$

The geometric explanation to this is just the Pythagorean theorem. Where we just solved for r. And since we are using norms, it can be expanded to higher dimensions.



||y||
||r||
range(A)
||y_{LS}||