Group 22 - HW04

April 5, 2021

Exercise 3.6 You are given

$$W = [\begin{array}{cccc} w_1 & 0 & ... & 0 \\ 0 & w_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & w_N \end{array}]$$

where, for all $n\epsilon\{1,..,N\}$, we have $w_n > 0$. Show that

$$< x, z >_w = x^T W z$$

is an inner product on R^N. What happens if, for all $n\epsilon\{1,..,N\}$, we have instead $w_n > 0$?

Solution We need to prove the following things:

- $< x, x >_w \geq 0, \forall x\epsilon R^N$
- $< x, x >_w = 0 \iff x = 0$
- $< x, z >_w = < z, x >_w$
- $< \kappa z + \lambda y, x >_w = \kappa < z, x >_w + \lambda < y, x >_w$
- We define, $X = \ [x_1 \quad .. \quad x_N]$

$$< X, X >_W = X^T W X = \ [x_1 \quad .. \quad x_N] \ [\begin{array}{cccc} w1 & 0 & ... & 0 \\ 0 & w2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & wN \end{array}] \ [x_1 \quad .. \quad x_N]^T \ =$$

$$= \ [x_1 w_1 \quad .. \quad x_N w_N] \ \ [x_1 \quad .. \quad x_N]^T \ = x_1^2 * w_1 + .. + x_N^2 w_N = \sum_{i=1}^{N} x_i^2 w_i \geq 0, \forall x \epsilon R^N$$

- For the second property,

$$< X, X >_W = .. = \sum_{i=1}^{N} x_i^2 w_i = 0, \forall x = 0$$

- For the third property, we will use a property of the dot product ($a * b = ab^T$)

  $< X, Z >_W = X^T W Z = X * (WZ) = (WZ) * X = (WZ)^T X = Z^T W^T X$

  but, $W^T = W$ so,

  $$< X, Z >_W = Z^T W X =< Z, X >_W$$

- For the fourth property,

  $< \kappa z + \lambda y, x >_w = (\kappa Z + \lambda Y)^T W X = (\kappa Z^T + \lambda Y^T) W X = \kappa Z^T W X + \lambda Y^T W X =$

  $$\kappa < Z, X >_w + \lambda < Y, X >_w$$

  So, given that all properties are proven, $< x, z >_w = x^T W z$ is an inner product on R^N.

Now if for all $n \epsilon \{1, .., N\}$, we have instead $w_n > 0$, then the second property would not fulfill, since it would be valid to say that

$$< x, z >_w = 0 \Leftarrow x = 0$$

but it would be wrong to say that

$$< x, z >_w = 0 \Rightarrow x = 0$$

So, if for all $n \epsilon \{1, .., N\}$, we have instead $w_n > 0$, then the $< x, z >_w = x^T W z$ is NOT an inner product on R^N

Exercise 3.7 Continuing from the above, show that,

$$< X, Z >_W = X^T W Z$$

is a norm on R^N. What happens if, for all $n \epsilon \{1, .., N\}$, we have instead $w_n > 0$?

Solution We need to prove the following things:

- $< x, x > \geq 0, \forall x \epsilon R^N$
- $< x, x >= 0 \iff x = 0$
- $< kx, kx >= |k| < x, x >$
- $||U + V||_W^2 \leq (||U||_W^2 + ||V||_W^2)^2$
- For the first property,

  $$||X||_W^2 =< X, X >_W = X^T W X = .. = \sum_{i=1}^{N} x_i^2 w_i \geq 0, \forall x \epsilon R^N$$

- For the second property,

$$||X||_W^2 =< X, X >_W= X^T W X = .. = \sum_{i=1}^{N} x_i^2 w_i = 0, \forall x = 0$$

- For the trird property,

$$||kX||_W = (kX)^T W(kX) = .. = \sum_{i=1}^{N} (kx_i)^2 w_i = \sum_{i=1}^{N} (kx_i)^2 \sum_{i=1}^{N} w_i =$$

$$\sqrt{k^2 (\sum_{i=1}^{N} w_i)||X||^2} = |k|||X||_W$$

- For the fourth property,

$$||U + V||_W^2 = (U + V)^T W(U + V) = (U^T + V^T)W(U + V) =$$

$$= U^T WU + U^T WV + V^T WU + V^T WV$$

On the other hand,

$$(||U||_W^2 + ||V||_W^2)^2 = U^T WU + V^T WV + 2||U||||V||$$

So we need to show that $2||U||||V|| \geq V^T WU + U^T WV$

But

$$U^T WU + V^T WV = U^T WU + U^T WU = 2U^T WU$$

So we need to show that $||U||||V|| \geq U^T WV$which is true because of the Cauchy-Schwarz inequality.

Now if for all $n \epsilon \{1, .., N\}$, we have instead $w_n \geq 0$, then the second property would not fulfill since,$< x, z >_w= 0$ does not mean that $x = 0$.

Exercise 3.8 Suppose A is an M x N matrix (with M > N), b is an M x 1 vector and w is an M x 1 vector with strictly positive components. Find the matrix $\hat{A}$which minimizes

$$\sum_{m=1}^{M} w_m (\sum_{n=1}^{N} A_{mn} x_n - b_m)^2$$

with respect to A.

Solution Firstly,

$$\sum_{m=1}^{M} w_m(\sum_{n=1}^{N} A_{mn}x_n - b_m)^2 = \sum_{m=1}^{M} w_m(Ax-b)_m^2$$

We define the MxM table W that has diagonal elements, the elements of vector w_n

$$W = [\begin{array}{cccc} w_1 & 0 & ... & 0 \\ 0 & w_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & w_M \end{array}]$$

Then,

$$\sum_{m=1}^{M} w_m(Ax-b)_m^2 = ||Ax-b||_W^2$$

So, the same way with the norm $||u||_w$ from exercise 3.7,

$$J(A) = ||Ax-b||_W^2 = (Ax-b)^T W(Ax-b)$$

In order to minimize J, we will take the derivative,

$$\nabla_A J(A) = \nabla_A = \nabla_A (Ax-b)^T W(Ax-b) = \frac{\partial((Ax-b)^T W(Ax-b))}{\partial A}$$

$$\nabla_A J(A) = 2W(Ax-b)x^T$$

So, we can find the$\hat{A}$ that minimizes $J(A)$ from the equation,

$$J(\hat{A}) = 0 \Rightarrow 2W(\hat{A}x-b)x^T = 0 \Rightarrow W(\hat{A}x-b)x^T = 0$$

The table W is inversable, since $w_n > 0$ and it is a diagonal table, so,

$$det(W) = w_1 w_2 .. w_n > 0$$

So,

$$W(\hat{A}x-b)x^T = 0 \Rightarrow W^{-1}W(\hat{A}x-b)x^T = 0 \Rightarrow \hat{A}xx^T - bx^T = 0 \Rightarrow \hat{A}xx^T = bx^T$$

The multiplication $xx^T$is 1x1 and non-zero, so the $(xx^T)^{-1}$exists. So,

$$\hat{A} = bx^T(xx^T)^{-1}$$

Exercise 3.11 Suppose A is an MN matrix (with M > N) and b is an M1 vector. Consider the following iteration

$$x^{(0)} = 0$$

$$x^{(t+1)} = x^{(t)} - \delta A^T(Ax^{(t)} - b)$$

Show that:

- if $x^{(t+1)} = x^{(t)}$ then $x^{(t)} =\hat{x} = (A^T A)^{-1} A^T b$
- there exists some δ_0 (find it) such that $\delta\epsilon(0, \delta_o) \Rightarrow< \lim_{t->\infty} x^{(t)} = \hat{x}$

Solution

- When $x^{(t+1)} = x^{(t)}$then,

$$\delta A^T(Ax^{(t)} - b) = 0 \Rightarrow A^T(Ax^{(t)} - b) = 0 \Rightarrow A^T Ax^{(t)} - A^T b = 0 \Rightarrow$$
$$A^T Ax^{(t)} = A^T b \Rightarrow x^{(t)} = (A^T A)^{-1} A^T b$$

- For the second question,

$$x^{(t+1)} = x^{(t)} - \delta A^T(Ax^{(t)} - b) \Rightarrow$$
$$x^{(t+1)} = x^{(t)} - \delta A^T Ax^{(t)} + \delta A^T b \Rightarrow$$
$$x^{(t+1)} = x^{(t)}[I - \delta A^T A] + \delta A^T b \Rightarrow$$
$$\lim_{t->\infty} x^{(t+1)} = \lim_{t->\infty} x^{(t)}[I - \delta A^T A] + \lim_{t->\infty} \delta A^T b \Rightarrow$$
$$\lim_{t->\infty} x^{(t)} = \lim_{t->\infty} (x^{(t)})[I - \delta A^T A] + \delta A^T b \Rightarrow$$
$$\lim_{t->\infty} x^{(t)}[I - I + \delta A^T A] = \delta A^T b \Rightarrow$$
$$\lim_{t->\infty} x^{(t)}[\delta A^T A] = \delta A^T b$$

So, when δ is non-zero

$$\lim_{t->\infty} x^{(t)}[A^T A] = A^T b \Rightarrow$$
$$\lim_{t->\infty} x^{(t)} = (A^T A)^{-1} A^T b \Rightarrow$$
$$\lim_{t->\infty} x^{(t)} = \hat{x}, \forall \delta \epsilon R^*$$

So, δ should be greater than zero, $\delta > 0$.

Exercise 3.12 Suppose that we have 1 x N vectors $a_1 a_2 .. a_n$ and scalars $b_1 b_2 .. b_n$. Now, for every $m \epsilon N$ define

$$A^{(m)} = [a_1 a_2 .. a_m]$$
$$x = [x_1 x_2 .. x_N]$$
$$b^{(m)} = [b_1 b_2 .. b_n]$$

Let $\hat{x}$ be the vector that minimizes $||A^{(m)} x - b^{(m)}||^2$. Consider the following iterarions.

$$G^{(N)} = (A^{(N)})^T A^{(N)}, G^{(m+1)} = G^{(m)} + a_m^T a_m$$
$$h^{(N)} = (h^{(N)})^T b^{(N)}, h^{(m+1)} = h^{(m)} + a_m^T b_m$$
$$x^{(N)} = (G^{(N)})^{-1} h^{(N)}, x^{(m+1)} = (G^{(m)})^{-1} h^{(m)}$$

Show that, for every $m \epsilon N$, $\hat{x}$ minimizes $||A^{(m)} x - b^{(m)}||^2$. Give an interpretation

Solution From theory, we know that the $\hat{x}$ that minimizes $||Ax-b||^2$is

$$\hat{x}=(A^TA)^{-1}A^Tb$$

For $m=\{N,N+1\}$
$m=N$

$$G^{(N+1)}=G^{(N)}+a_{N+1}^Ta_N$$
$$h^{(N+1)}=h^{(N)}+a_{N+1}^Ta_N$$

$m=N+1$

$$G^{(N+2)}=G^{(N+1)}+a_{N+2}^Ta_{N+2}=G^{(N)}+a_{N+1}^Ta_{N+1}+a_{N+2}^Ta_{N+2}$$

$$h^{(N+2)}=h^{(N+1)}+a_{N+2}^Tb_{N+2}=h^{(N)}+a_{N+1}^Tb_{N+1}+a_{N+2}^Tb_{N+2}$$

For $m=k,\ k\epsilon[N,\infty]$

$$G^{(k+1)}=G^{(k)}+a_{k+1}^Ta_{k+1}=(A^{(N)})^TA^{(N)}+a_{N+1}^Ta_{N+1}+..+a_{k+1}^Ta_{k+1}$$

$$h^{(k+1)}=h^{(k)}+a_{k+1}^Ta_{k+1}=(A^{(N)})^Tb^{(N)}+a_{N+1}^Tb_{N+1}+..+a_{k+1}^Tb_{k+1}$$

But

$$(A^{(N)})^TA^{(N)}=a_1^Ta_1+..+a_N^Ta_N$$
$$(A^{(N)})^Tb^{(N)}=a_1^Tb_1+..+a_N^Tb_N$$

So,

$$G^{(m+1)}=a_1^Ta_1+..+a_{m+1}^Ta_{m+1}=(A^{(m+1)})^TA^{(m+1)}$$
$$h^{(m+1)}=a_1^Tb_1+..+a_{m+1}^Tb_{m+1}=(A^{(m+1)})^Tb^{(m+1)}$$

So,

$$x^{(m+1)}=(G^{(m+1)})^{-1}h^{(m+1)}\Rightarrow$$
$$x^{(m+1)}=[(A^{(m+1)})^TA^{(m+1)}]^{-1}(A^{(m+1)})^Tb^{(m+1)}$$

which shows that for every $m\epsilon N$, $\hat{x}$ minimizes $||A^{(m)}x-b^{(m)}||^2$.