Problem Set 4

Contents

Problem 1

Explain:

(a) why $A^T A$ is not singular when matrix A has independent columns;
(b) why A and $A^T A$ have the same nullspace

Solution

(a)

A has independent columns $\Rightarrow$ it is full-rank, because column rank equals row rank. Then it has an inverse matrix A^{-1} and:

$$\begin{aligned} A^T A &= M, \\ A^T A A^{-1} &= M A^{-1}, \\ A^T &= M A^{-1} \end{aligned}$$

Because A^T is a full-rank matrix ($rk(A^T) = rk(A)$), M also has to be a full-rank matrix, thus, not singular.

(b)

Nullspace is a space spanned by all vectors x such that $Ax = 0$. Let x be any vector of that space $N(A)$. Then:

$$\forall x \in N(A) \rightarrow A^T A x = A^T (Ax) = A^T 0 = 0.$$

We've proven that $N(A) \subset N(A^T A)$. To prove that $N(A^T A) \subset N(A)$, we do the following. Let $A^T A y = 0$, but $Ay \neq 0$. Then:

$$\begin{aligned} A^T A y = 0; \quad A y = b \neq 0, \\ A^T b = 0 \Rightarrow b \in N(A^T). \end{aligned}$$

But $N(A^T) \perp R(A^T) = C(A)$. This means that b is not from column space of A, which is a contradiction, as $b = Ay$. Thus $b = Ay = 0$ and $N(A^T A) \subset N(A)$. Uniting the results, we obtain $N(A) = N(A^T A)$.

Problem 2

A plane in $\mathbb{R}^3$ is given by the equation $x_1 - 2x_2 + x_3 = 0$.

(a) Identify two orthonormal vectors u_1 and u_2 that span the same plane.

(b) Find a projector matrix P that projects any vector from $\mathbb{R}^3$ to the plane and a projector $P_\perp$ that projects any vector to the direction normal to the plane.

(c) Using these two projectors find the unit normal to the plane and verify that it agrees with a normal found by calculus methods (that use the gradient).

Solution

(a)

We find any two lines belonging to the plane, vectorizethem, then perform orthogonalization and normalize the vectors.

The first vector $\hat{u}_1$ we can obtain by plugging $x_1 = 0$ into the plane equation. Then,

$$-2x_2 + x_3 = 0; \quad x_3 = 2x_2$$

and

$$\hat{u}_1 = \frac{1}{\sqrt{5}}[0 \quad 1 \quad 2]^T.$$

The second $\hat{u}_2$ we find by plugging $x2 = 0$:

$$x_1 + x_3 = 0; \quad x_1 = -x_3$$

and

$$\hat{u}_2 = \frac{1}{\sqrt{2}}[1 \quad 0 \quad -1]^T.$$

The orthogonalization:

$$u_2 = \hat{u}_2 - proj_{u_1}(u_2) = \hat{u}_2 - \hat{u}_1 \frac{(\hat{u}_1, \hat{u}_2)}{|\hat{u}_1||\hat{u}_1|}$$

By this operations we receive:

$$u_1 = \hat{u}_1,$$
$$u_2 = \frac{1}{\sqrt{30}}[5 \quad 2 \quad -1]^T$$

(b)

In our case, as u_1 and u_2 are orthonormal, plane projector P can be built as:

$$P = UU^T,$$
$$U = [u_1 \quad u_2]$$

We receive the following matrix:

$$P = \begin{bmatrix} \frac{5}{6} & \frac{1}{3} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{6} & \frac{1}{3} & \frac{5}{6} \end{bmatrix}$$

Now, to obtain $P_\perp$, which is basically nn^T, where n is a vector normal to the plane, we find the vector, whose coordinates are the coefficients in the equation of the plane:

> **Note**
>
> This is the gradient method

$$n = \frac{1}{\sqrt{6}}[1 \quad -2 \quad 1]$$

and put in into the matrix:

$$P_\perp = nn^T =$$
$$\begin{bmatrix} \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix}$$

(c)

> **Note**
>
> As long as I've done the task in the wrong sequence, here is projector $P_\perp$ obtained by other means: $P_\perp$ projects to the nullspace of the P. So we find the vector that spans the nullspace of P and make a projector out of it.

$$N(P) = [-1 \quad 2 \quad -1]^T$$

Well, from this point it is obvious, because the steps are the same.

Note

Below is the code for all necessary computations that were avoided being done by hand.

```
import sympy as sp

############ Finding u_2

u1h = sp.Matrix([0, 1, 2])*1/sp.sqrt(5)
u2h = sp.Matrix([1, 0, -1])*1/sp.sqrt(2)

u2 = u2h - u1h*u1h.dot(u2h)
u2 = u2.normalized()

u2
```

$$\begin{bmatrix} \frac{\sqrt{30}}{6} \\ \frac{\sqrt{30}}{15} \\ -\frac{\sqrt{30}}{30} \end{bmatrix}$$

```
############ Finding P

u1 = u1h

U = u1.row_join(u2)

P = U*U.T
P
```

$$\begin{bmatrix} \frac{5}{6} & \frac{1}{3} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{6} & \frac{1}{3} & \frac{5}{6} \end{bmatrix}$$

```
############ Finding P_perp

n = sp.Matrix([1, -2, 1]).normalized()

Pp = n*n.T
Pp
```

$$\begin{bmatrix} \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix}$$

```
############ P_perp by the means of the nullspace

a = sp.Matrix(P.nullspace())
a = a.normalized()
a*a.T
```

$$\begin{bmatrix} \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix}$$

Problem 3

Let $M = span\{v_1, v_2\}$, where $v_1 = \begin{bmatrix} 1 & 0 & 1 & 1 \end{bmatrix}^T$, $v_2 = \begin{bmatrix} 1 & -1 & 0 & -1 \end{bmatrix}^T$.

(a) Find the orthogonal projector P_M on M.

(b) Find the kernel (nullspace) and range (column space) of P_M.

(c) Find $x \in M$ which is closest, in 2-norm, to the vector $a = \begin{bmatrix} 1 & -1 & 1 & -1 \end{bmatrix}^T$.

Solution

(a)

```
v1 = sp.Matrix([1, 0, 1, 1])
v2 = sp.Matrix([1, -1, 0, -1])

v1.dot(v2)
```

$$0$$

$v_1 \perp v_2$, thus, we only have to normalize them and construct a projector as $P_M = QQ^T$:

```
########## Constructing matrix Q from v_1 and v_2

v1 = v1.normalized()
v2 = v2.normalized()

M = v1.row_join(v2)
M
```

$$\begin{bmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ 0 & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} & 0 \\ \frac{\sqrt{3}}{3} & -\frac{\sqrt{3}}{3} \end{bmatrix}$$

```
########### Constructing projector

Pm = M*M.T
Pm
```

$$\begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

(b)

Following the standard procedure of finding nullspace and column space (asking sympy to do this for us):

```
######## Null space:
N = Pm.nullspace()
N[0].row_join(N[1])
```

$$\begin{bmatrix} -1 & -1 \\ -1 & -2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

```
######### Column space:
C = Pm.columnspace()
C[0].row_join(C[1])
```

$$\begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

(c)

For this we need to project vector a onto the M. This projection would be the closest to a vector $x \in M$ in terms of 2-norm. So,

$$P_m a = b \in M$$

```
a = sp.Matrix([1, -1, 1, -1])
b = Pm*a
b
```

$$\begin{bmatrix} \frac{4}{3} \\ -1 \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$$

Let’s check the error norm:

```
(a - b).norm()
```

$$\frac{\sqrt{6}}{3}$$

Problem 4

(a)

Using the determinant test, find c and d that make the following matrices positive definite:

$$A = \begin{bmatrix} c & 1 & 1 \\ 1 & c & 1 \\ 1 & 1 & c \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & d & 4 \\ 3 & 4 & 5 \end{bmatrix}$$

(b) A positive definite matrix cannot have a zero (or a negative number) on its main diagonal. Show that the matrix

$$A = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 5 \end{bmatrix}$$

is not positive by finding x such shat $x^T Ax \leq 0$.

Solution

(a)

Both A and B are symmetric and quadratic. We can use Silvester’s criteria:

```
c, d = sp.symbols('c, d')
from sympy.solvers.inequalities import solve_poly_inequality
from sympy import Poly

A = sp.Matrix([
    [c, 1, 1],
    [1, c, 1],
    [1, 1, c]
])

B = sp.Matrix([
    [1, 2, 3],
    [2, d, 4],
    [3, 4, 5]
])

Inequalities_A = []
for i in range(0, max(A.shape)):
    det = A[0:i+1, 0:i+1].det()
    Inequalities_A.append(det)

Solutions_A = []

for ineq in Inequalities_A:
    Solutions_A.append(solve_poly_inequality(Poly(ineq), '>'))

Solutions_A
```

```
[[Interval.open(0, oo)],
 [Interval.open(-oo, -1), Interval.open(1, oo)],
 [Interval.open(-2, 1), Interval.open(1, oo)]]
```

We found the conditions for all minors to be positive. Now let's unite the results and find c:

```
Unions_A = [sp.sets.Union(*sol) for sol in Solutions_A]
Intersection_A = sp.sets.Intersection(*Unions_A)
Intersection_A
```

$$(1, \infty)$$

With any $c > 1$ matrix A is positive-definite.

Now let's do the same for matrix B:

```
Inequalities_B = []

for i in range(0, max(B.shape)):
    det = B[0:i+1, 0:i+1].det()
    Inequalities_B.append(det)

Inequalities_B
```

```
[1, d - 4, 12 - 4*d]
```

As can be proven with simple calculations by hand, we have non-intersecting sets of possible d. Thus, matrix B is not positive-definite with any d.

(b)

```
from sympy import abc
x1, x2, x3 = abc.symbols('x1, x2, x3')

x = sp.Matrix([x1, x2, x3])

A = sp.Matrix([
    [4, 1, 1],
    [1, 0, 2],
    [1, 2, 5]
])

b = x.T*A*x
b[0]
```

$$x_1 (4x_1 + x_2 + x_3) + x_2 (x_1 + 2x_3) + x_3 (x_1 + 2x_2 + 5x_3)$$

It is obvious that plugging $x_1 = 0; x_3 = 0$ into the result will give us 0 with any x_2. Thus proven.

Also, only x_2 has no quadratic term, thus, we can decrease the product by decreasing the term x_2 with fixed x_1 and x_3:

```
subs = {x1: 1, x2: -10, x3: 1}
b = b.subs(subs)
b
```

$$[-49]$$

Problem 5

Matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$ is positive definite. Explain why and determine the minimum value of $z = x^T Ax + 2b^T x + 1$, where $x = [x_1 \quad x_2 \quad x_3]^T$ and $b^T = [1 \quad -2 \quad 1]$

Solution

Warning

To be clear, I don't exactly understand the meaning "positive-definite" applied to a non-symmetric matrix. But okay...

```
A = sp.Matrix([
    [1, 1, 0],
    [2, 3, 1],
    [1, 1, 4]
])

for i in range(0, max(A.shape)):
    string = '{} minor is: '.format(i+1)
    print(string, A[0:i+1, 0:i+1].det())
```

```
1 minor is:  1
2 minor is:  1
3 minor is:  4
```

So it is positive (in some way, e.g. Totally Positive Matrix) due to positivity of its minors.

```
from sympy import functions
x1, x2, x3 = abc.symbols('x1, x2, x3')

x = sp.Matrix([x1, x2, x3])
b = sp.Matrix([1, -2, 1])

z = x.T*A*x + 2*b.T*x + 1*sp.Matrix([1])
f = z[0]
f_difs = []

for x_ in x:
    f_difs.append(f.diff(x_))

f_difs = sp.Matrix(f_difs)
sols = sp.solve(f_difs)
sols
```

```
{x1: -8, x2: 53/11, x3: -5/11}
```

We've got the x which corresponds to the minimal value of z. Let's find it:

```
z = z.subs(sols)
z[0]
```

$$-\frac{188}{11}$$

Note

This was the straightforward solution. Also we may perform some matrix manipulations first.

$$\begin{aligned} z = x^T A x + 2b^T x + 1 = x^T A_+ x + 2b^T x + 1 = \frac{1}{2} x^T (A + A^T) x + 2b^T x + 1; \\ \frac{dz}{dx} = (A + A^T)x + 2b = 0, \\ (A + A^T)x = -2b \quad \Rightarrow \quad x = -2(A + A^T)^{-1} b. \end{aligned}$$

From this point it is just a matter of calculations:

```
y = -2*(A + A.T).inv()*b
y
```

$$\begin{bmatrix} -8 \\ \frac{53}{11} \\ -\frac{5}{11} \end{bmatrix}$$

The same result.

Problem 6

Explain these inequalities from the definitions of the norms:

$$||ABx|| \leq ||A|| ||Bx|| \leq ||A|| ||B|| ||x||,$$

and deduce that $||AB|| \leq ||A|| ||B||$.

Solution

For a given x:

$$\begin{aligned} ||ABx||_2^2 = (ABx)^T (ABx) = x^T B^T A^T ABx, \\ ||Bx||_2^2 = (Bx)^T (Bx) = x^T B^T Bx, \\ ||x||_2^2 = x^T x, \\ ||A||_2^2 = \max_{x \in \mathbb{R}^n} \left\{ \frac{(Ax)^T (Ax)}{x^T x} \right\} = \max_{x \in \mathbb{R}^n} \left\{ \frac{x^T A^T Ax}{x^T x} \right\}, \\ ||B||_2^2 = \max_{x \in \mathbb{R}^n} \left\{ \frac{(Bx)^T (Bx)}{x^T x} \right\} = \max_{x \in \mathbb{R}^n} \left\{ \frac{x^T B^T Bx}{x^T x} \right\}. \end{aligned}$$

Thus, we may rewrite some inequalities:

$$||ABx|| = x^T B^T A^T ABx \frac{x^T B^T Bx}{x^T B^T Bx} = \frac{x^t B^T A^T ABx}{x^T B^T Bx} x^T B^T Bx =$$

$$= \frac{y^T A^T Ay}{y^T y} x^T B^T Bx \leq \max_{y \in \mathbb{R}^n} \left\{ \frac{y^T A^T Ay}{y^T y} \right\} ||Bx|| = ||A|| ||Bx||;$$

$$||Bx|| = x^T B^T Bx = x^T B^T Bx \frac{x^T x}{x^T x} = \frac{x^T B^T Bx}{x^T x} x^T x \leq \max_{z \in \mathbb{R}^n} \left\{ \frac{z^T B^T Bz}{z^T z} \right\} x^T x = ||B|| ||x||$$

So, we've proven that $||ABx|| \leq ||A|| ||Bx||$ and $||Bx|| \leq ||B|| ||x||$,
$\Rightarrow ||ABx|| \leq ||A|| ||Bx|| \leq ||A|| ||B|| ||x||$.

The first line we can use to prove that

$$||AB|| = \max_{x \in \mathbb{R}^n} \frac{||ABx||}{||x||} = \max_{x \in \mathbb{R}^n} \frac{||ABx||}{||x||} \frac{||Bx||}{||Bx||} =$$

$$\max_{x \in \mathbb{R}^n} \frac{||ABx||}{||Bx||} \frac{||Bx||}{||x||} \leq \max_{x \in \mathbb{R}^n} \frac{||ABx||}{||Bx||} \max_{x \in \mathbb{R}^n} \frac{||Bx||}{||x||} = ||A|| ||B||.$$

Print to PDF

Problem 7

Compute by hand the norms and condition numbers of the following matrices:

$$A_1 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

Solution

A_1 is symmetric, in this case its norm is the largest eigenvalue. Thus,

$$\begin{aligned} (2-\lambda)^2 - 1 = 0, \\ \lambda_1 = 3, \quad \lambda_2 = 1, \\ ||A_1|| = \max_i \left\{\lambda_i\right\} = 3, \\ \kappa_1 = \frac{\max_i \left\{\lambda_i\right\}}{\min_i \left\{\lambda_i\right\}} = 3 \end{aligned}$$

A_2 is skew-symmetric, for non-symmetric matrices the norm equals to the greatest singular value:

$$\begin{aligned} A_2^T A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \\ \lambda_1 = \lambda_2 = 2, \\ \sigma_1 = \sigma_2 = \sqrt{2} \end{aligned}$$

```
A2 = sp.Matrix([
    [1, 1],
    [-1, 1]
])

A2.singular_value_decomposition()[1]
```

$$\begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$$

$$\begin{aligned} \sigma_1 = \sigma_2 = \sqrt{2}, \\ ||A_2|| = \max_i \left\{\sigma_i\right\} = \sqrt{2}, \\ \kappa_2 = \frac{\max_i \left\{\sigma_i\right\}}{\min_i \left\{\sigma_i\right\}} = 1. \end{aligned}$$

$$

By Arseniy Buchnev