

Applied Advanced Optimisation iRAT 1

Zhao, Amor (CID: 02019680)

January 16, 2025

Question 1

State all the languages you speak.

English, Chinese (Mandarin).

Question 2

Write the solution of $\min_x \|Ax - b\|_2^2$.

To minimise $(Ax - b)^T(Ax - b) = x^T A^T A x - 2x^T A^T b + b^T b$,

we take the derivative with respect to x and set it to zero: $2A^T A x - 2A^T b = 0$.

Thus $x^* = (A^T A)^{-1} A^T b$.

Question 3

Explain why in the norm approximation problem, the ℓ_1 -norm generates a large number of zero residuals.

Because the ℓ_1 -norm adds the absolute values together, when minimising the residues using ℓ_1 -norm, we get a very sparse solution.

Question 4

Which penalty function would you pick to reduce sensitivity to outliers?

ℓ_1 -norm or the Huber penalty function.

Question 5

Give the solution to the ℓ_2 least-norm problem $\min_x \|x\|_2^2$ s.t. $Ax = b$.

$x^* = A^\top (AA^\top)^{-1} b$.

Question 6

Give the solution of the Tikhonov regularisation problem.

We have the general equation of Tikhonov regularisation problem:

$$\min_x ||Ax - b||_2^2 + \gamma ||x||_2^2$$

Objective function:

$$J(x) = ||Ax - b||_2^2 + \gamma ||x||_2^2$$

We take the derivative with respect to x and set it to zero:

$$\nabla_x J(x) = 2A^T Ax - 2A^T b + 2\gamma x = 0$$

$$A^T Ax + \gamma x = A^T b$$

Thus:

$$x = (A^T A + \gamma I)^{-1} A^T b$$

Question 7

Make a comparison between the solution of the nominal least-square, stochastic least-square and worst-case least-square.

The nominal least-square solution achieves best result when $u = 0$, ie. when there are no significant outliers.

The stochastic least-square solution performs better than nominal least-square solution with large u , ie. when the data is noisy.

The worst-Case least-square is least sensitive to large number of outliers, it has a similar performance across all u .

Question 8

Declare a variable x of dimension n in CVXPY.

According to <https://www.cvxpy.org/tutorial/intro/index.html>, we can declare:

```
1 import cvxpy as cp
2 x = cp.Variable(n)
```

Question 9

Declare the constraints $x + y = 1$ and $x - y \geq 1$ in CVXPY.

```
1     y = cp.Variable()
2     constraints = [
3         x + y == 1,
4         x - y >= 1
5     ]
```

Question 10

Which atomic function would you use in CVXPY to compute the ℓ_∞ -norm of a variable x ?

```
1     norm_inf_x = cp.norm(x, "inf")
```

Feedback

Your grade for the first iRAT should be visible on blackboard. I received 38 submissions. The most common mistake was on question 3. The majority of submissions has either no or one error. One submission was late, thus the standard policy was applied and the mark was capped to pass (50). Please, do not submit late. Model answers below. If an answer is missing below, it means that no one got it wrong.

2. Some submissions did not contain the answer, namely $x = \text{inverse}(A'A)A'b$

3. The l1-norm generates a large number of zero residuals because it has a constant slope. So no matter how small the residual is, there is a constant push to make it smaller. This is not the case for instance with the l2-norm. The l2-norm has a slope which is linear in the residual. As the residual approaches zero, the slope becomes zero and stops pushing. (Some replied that the l1 penalises more small residuals than big. This is wrong. The penalty is constant, so independent of how big or small the residual is. This is why the residuals are carried all the way to zero).

4. For instance Huber, or l1.

5. Some submissions did not contain the answer, namely $x = A' \text{inverse}(AA')b$

8.

```
1  import cvxpy as cp (I considered your answer correct also
   without this line)
2  x = cp.Variable(n)
```

9.

```
1  constraints = [x+y==1, x-y>=1]
```

10.

```
1  import cvxpy as cp (considered correct also without this
   line and the cp. below)
2  cp.norm(x,"inf") (considered correct also norm_inf)
```