Computational Methods for Detection, Estimation and Identification 2022/2023

Assignment 4

Manuel Alberto Dionísio dos Santos
2019231352

1 Overview

This assignment focus on rank deficient and ill-conditioned systems of equations and on the exploration of Singular Value Decomposition related methods to achieve solutions of this systems. As a case of study, we are provided with an arbitrary polynomial of degree 8 (1).

$$y = a\theta^8 + b\theta^7 + c\theta^6 + d\theta^5 + e\theta^4 + f\theta^3 + g\theta^2 + h\theta \qquad (1)$$

Apart from the polynomial to study, a function SimulationPoly8() is also provided. This function simulates readings of data pairs (y_i , θ_i), corresponding to points belonging to the 8^{th} degree polynomial, when provided with the number of data points to generate, the standard deviation of the additive Gaussian noise and the percentage of outliers in the data set.

The objective is to study the fitting to the described 8^{th} degree when considering values of 10, 16 and 32 for the number of data points to generate and 0.0015, 0.15 and 0.3 for the standard deviation of the additive Gaussian noise.

In order to more easily display the results on this report, unless explicitly said otherwise, the results will refer to the case in which the number of data points is 16 and $\sigma = 0.15$.

2 Point 1

The first task is to obtain a Least Square solution for the model. In order to achieve this task, the first step was to get the G matrix for the specific problem (2).

$$G = \begin{bmatrix} \theta_1^8 & \theta_1^7 & \theta_1^6 & \theta_1^5 & \theta_1^4 & \theta_1^3 & \theta_1^2 & \theta_1 \\ \theta_2^8 & \theta_2^7 & \theta_2^6 & \theta_2^5 & \theta_2^4 & \theta_2^3 & \theta_2^2 & \theta_2 \\ & & \vdots & & & \vdots & & \\ \theta_m^8 & \theta_m^7 & \theta_m^6 & \theta_m^5 & \theta_m^4 & \theta_m^3 & \theta_m^2 & \theta_m \end{bmatrix}_{m \times 8} \qquad (2)$$

Similarly to previous assignments, the Least Squares solution [2] is obtained using the Pseudo-Inverse of Moore-Penrose.

3 Point 2

The second task is to obtain the generalised inverse solution using the Singular Value Decomposition (SVD).

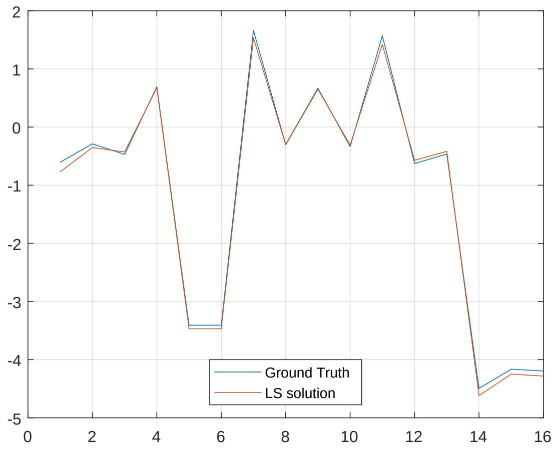

Figure 1: Least Squares solution for m = 16 and $\sigma = 0.15s$

Before starting, let's review some concepts. Singular Value Decomposition is used to factorize matrices into 3 components: U , S and V. In our case, the objective is to factorize G (3).

$$G = USV^T \qquad (3)$$

In this decomposition, S is a p by p diagonal matrix composed of the positive decreasing singular values of G, with p = rank(G); U denotes an m by m orthogonal matrix with columns that are unit basis vectors spanning the data space, R^m; and finally V denotes an n by n orthogonal matrix with columns that are unit basis vectors spanning the model space, R^n .

Since some of the singular values in S can be 0, the last m - p columns of U and the last n - p columns of V are multiplied by zeros, and as so we can simplify the SVD of G into its compact form (4).

$$G = U_p S_p V_p^T \qquad (4)$$

The SVD can then be used to compute a generalized inverse of G (6).

$$G^\dagger = V_p S_p^{-1} U_p^T \qquad (5)$$

Using the generalized inverse (6), the pseudoinverse solution can be defined (**??**).

$$m_\dagger = G^\dagger d = V_p S_p^{-1} U_p^T d \qquad (6)$$

Among the desirable properties of is that $G^\dagger$, and hence $m^\dagger$, always exist, in contrast to the Least Squares solution, that only exists when G is full rank.

When dealing with the generalized inverse, it's possible to find 4 separate cases, depending on the relations between n, m and p. In this assignment, the conditions n = p and p ¡ m are always true. As so, let's explore the implications of this situation.

$$Gm_\dagger = U_p S_p V_p^T S_p^{-1} U_p^T d = U_p U_p^T d \qquad (7)$$

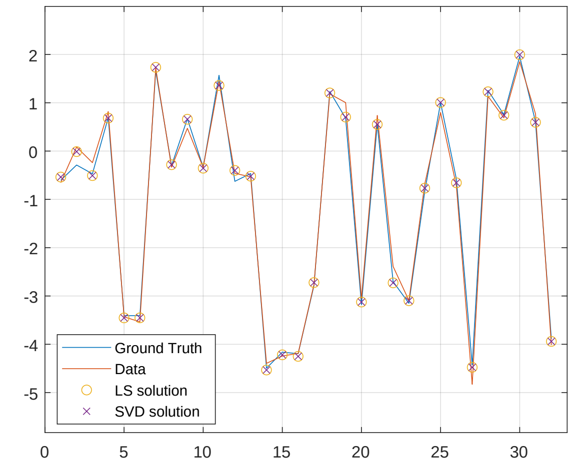

Figure 2: m = 16 and $\sigma = 0.15s$

The product $U_p U_p^T d$ gives the projection of d onto R(G). Thus $Gm_\dagger$ is the point in R(G) that is closest to d, and $m^\dagger$ is a least squares solution to Gm = d.

Having this conclusion in mind, let's compute the generalised inverse solution. Using the Matlab svd() function, and the equations obtained, the results obtained are shown in

As expected, solutions obtained for the SVD are the same as the ones obtained for the LS.

4 Point 3

To compute the confidence intervals, it's necessary to get the covariance matrix and, as so, the constant data element standard deviation. Note that the use of the constant data element standard deviation implies the calculation of the inverse cumulative distribution for a Student's t distribution.

For the calculation of s (constant data element standard deviation), equation (8) is used.

$$s = \frac{||residuals||_2}{\sqrt{v}} \tag{8}$$

Knowing s, and assuming $\sigma \approx s$, the covariance matrix of $m^\dagger$ is given by (9).

$$\begin{aligned} Cov(m_\dagger) &= G^\dagger Cov(d)(G^\dagger)^T \\ &= \sigma^2 G^\dagger (G^\dagger)^T \\ &= \sigma^2 V_p S_p^{-2} V_p^T \\ &= \sigma^2 \sum_{i=1}^{p} \frac{V_{.,i} V_{.,i}^T}{s_i^2} \end{aligned} \tag{9}$$

The last part needed is the inverse cumulative distribution for a Student's t distribution. Knowing that the Student's t distribution is symmetric, this issue can be resolved by running icd = tinv(0.975,degrees of freedom).

Finally, the confidence intervals of the solution are given by

$$m_\dagger \pm icd \cdot diag(Cov(m_\dagger))^{1/2} \tag{10}$$

The results can be seen of figure (3),(4),(5) and (6).

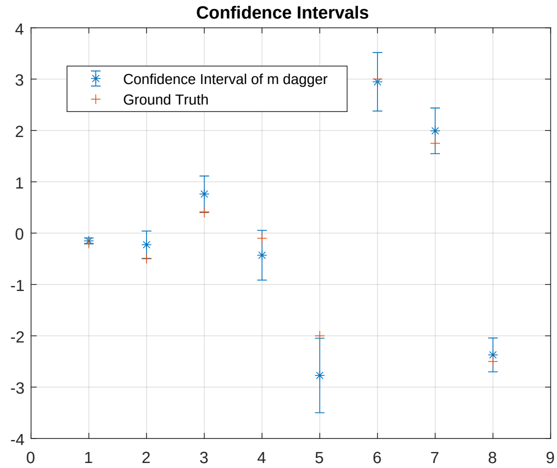

Figure 3: m = 16 and $\sigma = 0.15s$

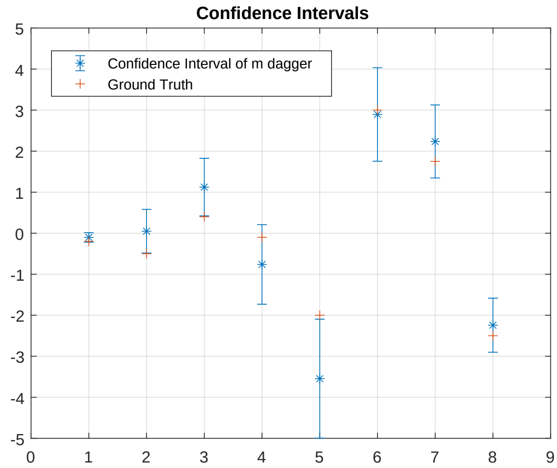

Figure 4: m = 16 and $\sigma = 0.3s$

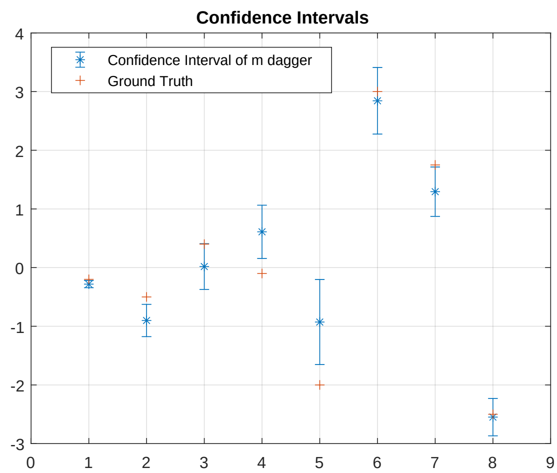

Figure 5: m = 32 and $\sigma = 0.15s$

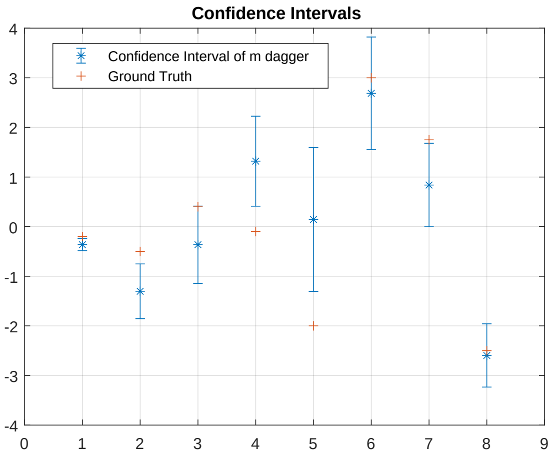

Figure 6: m $= 32$ and $\sigma = 0.3s$

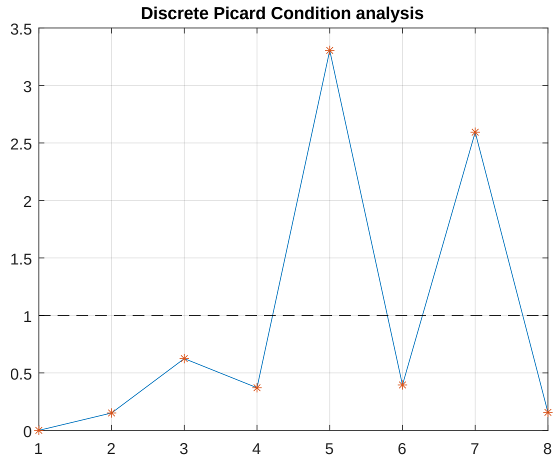

Figure 7: Discrete Picard condition

There are 2 main aspects to discuss:

- Due to the presence of the Student's distribution and it's properties for when the degrees of freedom are less or equal to 2, in the case in with m = 10, it's impossible to calculate the confidence intervals;
- It's possible to verify that the uncertainty of the results decreases with the increase in the number of data points (m) and with the decrease in the standard deviation;

5 Point 4

The condition number of a SVD solution is a measure of the sensitivity of the solution to changes in the input data. It is defined as the ratio of the largest singular value to the smallest singular value, as showed on equation (11). In practice, the condition number of the SVD solution can be used to determine whether the input data is ill-conditioned (i.e., highly sensitive to changes).

$$cond(G) = \frac{s_1}{s_k} \tag{11}$$

In our case, and for all the possible combinations of m and σ defined, the condition numbers obtained all revolve around 10^4. This is a high value, and as so suggests that the matrix is ill-conditioned, which means that small changes in the input data can lead to large changes in the output. Note that when m, the number of data points rises, the condition number decreases.

6 Point 5

A condition that insures solution stability is the discrete Picard condition. The discrete Picard condition is satisfied when the dot products of the columns of U and the data vector decay to zero more quickly than the singular values. Under this condition, we should not see instability due to small singular values.

Plotting the ratios described by the discrete Picard condition produces the results shown on figure (7).

Given the result, the best way to obtain the Truncated SVD solution is to truncate the singular values and vectors at index i = 4.

7 Point 6

The model resolution characterizes the linear relationship between $m^\dagger$ and m_true for a linear inverse problem when solved using the generalized inverse. It is given by equation (12).

$$\begin{aligned} R_m &= G^\dagger G \\ &= V_p S_p^{-1} U_p^T U_p S_p V_p^T \\ &= V_p V_p^T \end{aligned} \tag{12}$$

Figures (8) and (9) show the model resolution for the SVD solution and the Truncated SVD solution, respectively.

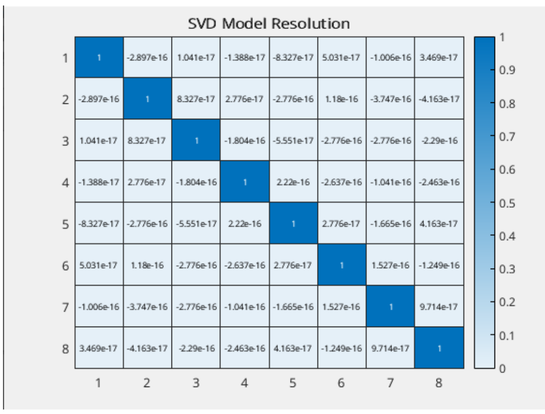

Figure 8: SDV solution model resolution

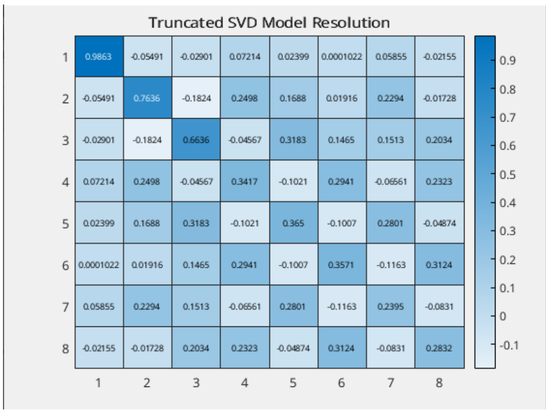

Figure 9: Truncated SDV solution model resolution

If N (G) is trivial, then rank(G) = p = n, and R^m is the n by n identity matrix. In this case the original model is recovered exactly and we say that the model resolution is perfect, as seen in (8). If N (G) is a nontrivial subspace of R^n , then p = rank(G) < n, so that R^m is a non-identity symmetric and singular matrix.

8 Point 7

The SVD solution is not very robust in itself, due to the presence of the smalles singular values. As so, a way to increase the robustness is to use the Truncated SVD solution, as seen on figure (10). Truncated SVD can be more robust than the full SVD in the presence of noise or missing data. The reason for this is that the full SVD considers all the singular values of a matrix, including those that may be small or close to zero, while the truncated SVD only considers the largest singular values.

In situations where the input data is noisy or contains missing values, the small or close-to-zero singular values in the full SVD can be particularly sensitive to errors, leading to inaccurate solutions. In contrast, the truncated SVD only considers the most significant singular values, which are less sensitive to noise and missing data, and thus the truncated SVD solution may be more robust in such situations.

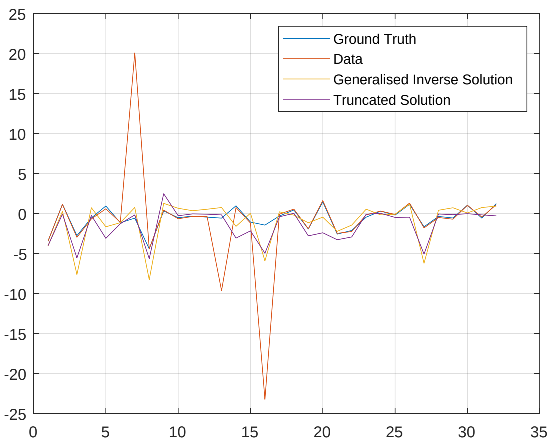

Figure 10: Robustness Comparison

References

[1] R. C. Aster, B. Borchers, and C. H. Thurber, Parameter Estimation and Inverse Problems. Elsevier, 2018.

[2] Nuno Gonçalves, sebenta_TCDEI_chapter1_linear_inverse_problems, University of Coimbra, 2023 [3] https://link.springer.com/article/10.10

[4] https://math.stackexchange.com/questions/500244/how-singular-is-a-matrix