DISCRETE INVERSE PROBLEMS
(**LAB - 2**)

Maris Ioannis

2022

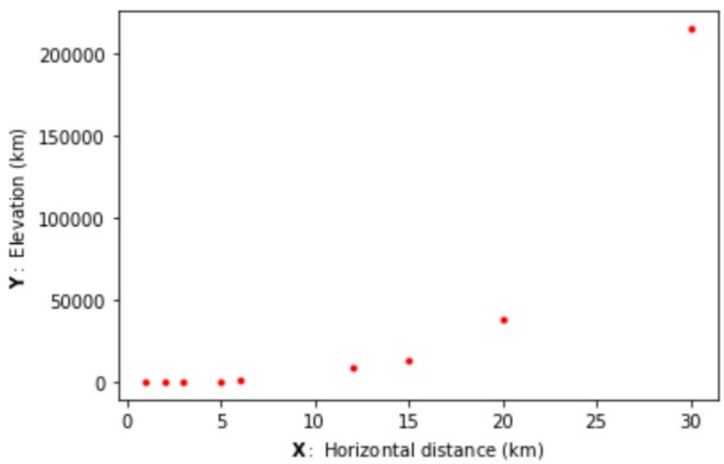
A spaceship was launched in space, its orbit was measured by monitoring station, with observations giving the position of the orbit as pairs of distance from the station on the ground surface and elevation. A crew member claimed that the spacecraft's orbit follows a higher polynomial curve in plane (**deg**$\rightarrow 3$).
The goal is to test the claim by studying its trajectory spaceship. In this context you should calculate the parameters of the trajectory with the member's hypothesis and find the data analysis matrices and parameters as well as the unit covariance matrices for the following **data**:

Horizontal distance (km)	Elevation (km)
1	23
2	49
3	107
5	442
6	905
12	8700
15	13200
20	38440
30	215300

Ignore the curvature of the earth despite the hints, Use SVD analysis, to determine the solution of the problem.

GITHUB LINK ← (clickable)

First of all, lets visualize our data:



Model:

$$\mathbf{d} = \mathbf{m_1} + \mathbf{m_2 x} + \mathbf{m_3 x^2} + \mathbf{m_4 x^3},$$

data:

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix}, \quad N = 9$$

parameters:

$$\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_M \end{bmatrix}, \quad M = 4$$

Hence, we get the following (linear) system:

$$\begin{aligned}
d_1 &= m_1 + m_2 + m_3 + m_4 \\
d_2 &= m_1 + 2m_2 + 4m_3 + 8m_4 \\
d_3 &= m_1 + 3m_2 + 9m_3 + 27m_4 \\
d_4 &= m_1 + 5m_2 + 25m_3 + 125m_4 \\
d_5 &= m_1 + 6m_2 + 36m_3 + 216m_4 \\
d_6 &= m_1 + 12m_2 + 144m_3 + 1728m_4 \\
d_7 &= m_1 + 15m_2 + 225m_3 + 3375m_4 \\
d_8 &= m_1 + 20m_2 + 400m_3 + 8000m_4 \\
d_9 &= m_1 + 30m_2 + 900m_3 + 27000m_4
\end{aligned}$$

$$\rightarrow \mathbf{d} = \begin{bmatrix} 23 \\ 49 \\ \vdots \\ 215300 \end{bmatrix},$$

We now write the linear system in the form of matrix equation, namely:

$$\rightarrow \mathbf{Gm} = \mathbf{d}, \quad \textbf{(Explicit Linear Form)}$$

$\mathbf{m} \in \mathbb{R}^4$ and $\mathbf{d} \in \mathbb{R}^9$, therefore, we get an **overdetermined** inverse problem, which we will solve it by using least squares method.

The matrix $\mathbf{G}$ is called the **data kernel** and is defined by the system structure:

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 5 & 25 & 125 \\ 1 & 6 & 36 & 216 \\ 1 & 12 & 144 & 1728 \\ 1 & 15 & 225 & 3375 \\ 1 & 20 & 400 & 8000 \\ 1 & 30 & 900 & 27000 \end{bmatrix},$$

(**"Normal Equations"**)

The measurements ($\mathbf{d_i}$) are equivalent, and in terms of their accuracy they follow a normal distribution, hence, we minimize the error :

$$\rightarrow \mathbf{E} = \mathbf{e^T e} = \sum_{i=1}^{N} e_i^2, \qquad \mathbf{e} = \mathbf{d} - \mathbf{d^{pre}},$$

If we solve the problem and get the solutions $\mathbf{m^{est}}$ the above equation (ELF) gives:

$$\mathbf{d^{pre}} = \mathbf{Gm^{est}}$$

and the total error $\mathbf{E}$ is:

$$E = \mathbf{e^T e} = (\mathbf{d} - \mathbf{d^{pre}})^T(\mathbf{d} - \mathbf{d^{pre}}) = (\mathbf{d} - \mathbf{Gm^{est}})^T(\mathbf{d} - \mathbf{Gm^{est}}) =$$

$$\sum_{i=1}^{N} \left[d_i - \sum_{j=1}^{M} \mathbf{G}_{ij} m_j^{est} \right] \left[d_i - \sum_{j=1}^{M} \mathbf{G}_{ij} m_j^{est} \right]$$

So, we differentiate for each parameter $m_q, q = 1, \dots, M$ and we ask this derivative to be zero, in order to minimize total error. From Matrix Calculus theory, we get:

$$\frac{\partial \mathbf{Au}}{\partial x} = \frac{\partial \mathbf{u}}{\partial x} \mathbf{A}^T.$$

We fix some m_i and differentiate the total error:

$$\frac{\partial \mathbf{E}}{\partial m_i} = \frac{\partial}{\partial m_i}\left(\sum_{i=1}^{N}\left[d_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right]\left[d_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right]\right) =$$

$$= \sum_{i=1}^{N}\frac{\partial}{\partial m_i}\left(\left[d_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right]\left[d_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right]\right) =$$

$$= \sum_{i=1}^{N}\frac{\partial}{\partial m_i}\left(\left[\mathbf{G}m_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right]\left[\mathbf{G}m_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right]\right) =$$

$$\sum_{i=1}^{N}\left(\mathbf{G}^T\left[\mathbf{G}m_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right] + \mathbf{G}^T\left[\mathbf{G}m_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right]\right) =$$

$$\sum_{i=1}^{N}\mathbf{G}^T\left(d_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est} + d_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right) = \sum_{i=1}^{N}\mathbf{G}^T\left(2d_i - 2\sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right) =$$

$$= 2\mathbf{G}^T\sum_{i=1}^{N}\left(d_i - \sum_{j=1}^{M}\mathbf{G}_{ij}m_j^{est}\right) = -2\mathbf{G}^T\big(\mathbf{G}\mathbf{m}^{\mathbf{est}} - \mathbf{d}\big) = 0 \iff$$

$$\iff \boxed{\mathbf{G^T G m^{est}} - \mathbf{G^T d} = 0.}$$

Hence,

$$\rightarrow \mathbf{m^{est}} = \big(\mathbf{G^T G}\big)^{-1}\mathbf{G^T d},$$

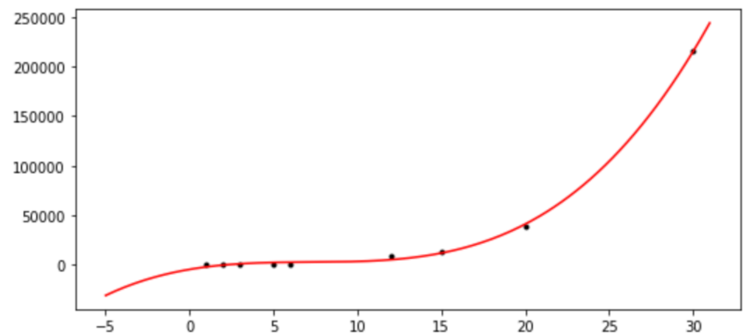
$$\mathbf{G^T G} = \begin{bmatrix} 9 & 94 & 1744 & 40480 \\ 94 & 1744 & 40480 & 1043380 \\ 1744 & 40480 & 1043380 & 28519384 \\ 40480 & 1043380 & 28519384 & 807439684 \end{bmatrix},$$

Therefore, we get the following solutions for the parameters,

$$\rightarrow \mathbf{m^{est}} = \begin{bmatrix} -4526.12816835896 \\ 2881.0401528196 \\ -385.877006164702 \\ 17.787867537097 \end{bmatrix},$$

$$\mathbf{d^{pre}} = \mathbf{G m^{est}} = \begin{bmatrix} -2013.17715416696 \\ -165.25294708179 \\ 1124.37165811914 \\ 2455.63088375861 \\ 2710.71991464232 \\ 5217.49988186275 \\ 11901.2006745794 \\ 41046.8127189282 \\ 214888.194369616 \end{bmatrix} \Rightarrow \|\mathbf{d^{obs}} - \mathbf{d^{pre}}\|_2 = 5772.525.$$

$$\rightarrow \mathbf{d} = -\,4526.128 + 2881.04x + -385.877x^2 + 17.788x^3,$$



Let's now solve the problem by using the **SVD** factorization (Singular Value Decomposition).

Singular value decomposition (SVD) of $m \times n$ matrix **G** has form:

$$\rightarrow \mathbf{G} = U\Sigma V^T$$

where U is $m \times m$ orthogonal matrix, V is $n \times n$ orthogonal matrix, and Σ is $m \times n$ diagonal matrix, with:

$$\Sigma_{ij} = \begin{cases} 0, & i \neq j \\ \sigma_i, & i = j \end{cases},$$

Diagonal entries σ_i, called **singular values** of **G**, are usually ordered such that $\sigma_1 > \sigma_2 > \cdots > \sigma_n$.
The columns u_i of U and v_i of V are called left and right **singular vectors**.

The vector

$$\mathbf{m}^{\mathbf{est}} = \sum_{\sigma_i \neq 0} \frac{u_i^T \mathbf{d}}{\sigma_i} v_i$$

minimizes the norm:

$$\|\mathbf{d} - \mathbf{Gm}\|,$$

if rank$(\mathbf{G}) = n$, then the solution is unique.
For ill-conditioned or rank deficient problems, "small" singular values can be omitted from summation to stabilize solution.

SVD can also handle the rank deficient case, if there are only k singular values $\sigma_j > \epsilon$ then take only the first k contributions.

The **natural generalized inverse** (or pseudoinverse) of general real $m \times n$ matrix **G** is given by:

$$\rightarrow \mathbf{G}^{-\mathbf{g}} = V_p \Sigma_p^{-1} U_p^T,$$

where U_p, V_p consist of the first p columns of U , V and Σ_p is the block diagonal matrix with all the non-zero singular values in the diagonal.

Hence, **minimum-norm** solution (**natural solution**) is given by:

$$\rightarrow \mathbf{m}^{\mathbf{est}} = \mathbf{G}^{-\mathbf{g}}\mathbf{d}.$$

In our case, we can compute the SVD analysis numerically, after calculations, we get the factorization:

$$\mathbf{G} = U\Sigma V^T,$$

where

$$U \approx \begin{bmatrix} 0 & 0.0056 & 0.2003 & 0.7098 & -0.0339 & -0.101 & -0.271 & -0.562 & -0.2352 \\ -0.0003 & 0.0205 & 0.3145 & 0.4374 & -0.0241 & 0.1523 & 0.1798 & 0.2277 & 0.7755 \\ -0.001 & 0.0436 & 0.3997 & 0.2105 & 0.0853 & 0.387 & 0.4396 & 0.3573 & -0.5622 \\ -0.0044 & 0.1099 & 0.4923 & -0.1191 & -0.5937 & -0.4916 & -0.2098 & 0.2905 & -0.0939 \\ -0.0076 & 0.151 & 0.5042 & -0.2281 & 0.7505 & -0.2609 & -0.1898 & -0.0342 & 0.0488 \\ -0.0609 & 0.4377 & 0.2391 & -0.2958 & -0.2163 & 0.6513 & -0.343 & -0.2554 & 0.0782 \\ -0.1189 & 0.5571 & -0.0073 & -0.0926 & -0.0892 & -0.2745 & 0.6503 & -0.3968 & 0.0573 \\ -0.2817 & 0.6173 & -0.3734 & 0.3077 & 0.1419 & -0.0844 & -0.2824 & 0.4383 & -0.0807 \\ -0.9501 & -0.2826 & 0.0894 & -0.0587 & -0.0204 & 0.0216 & 0.0263 & -0.0654 & 0.0121 \end{bmatrix},$$

$$V^T \approx \begin{bmatrix} -5.0 \cdot 10^{-5} & -0.001292 & -0.0353 & -0.999376 \\ 0.008703 & 0.100195 & 0.994305 & -0.035251 \\ 0.31918 & 0.942642 & -0.097704 & 0.002217 \\ 0.947654 & -0.318412 & 0.023774 & -0.000476 \end{bmatrix},$$

$$\Sigma \approx \begin{bmatrix} 28433.22714525 & 0 & 0 & 0 \\ 0 & 190.7258432612 & 0 & 0 \\ 0 & 0 & 5.823603928 & 0 \\ 0 & 0 & 0 & 0.91927311741 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \Sigma_p = \begin{bmatrix} 28433.22714525 & 0 & 0 & 0 \\ 0 & 190.7258432612 & 0 & 0 \\ 0 & 0 & 5.823603928 & 0 \\ 0 & 0 & 0 & 0.91927311741 \end{bmatrix} \Rightarrow$$

$$\mathbf{G}^{-\mathbf{g}} \approx \begin{bmatrix} 0.7427 & 0.4682 & 0.2389 & -0.0958 & -0.2075 & -0.2919 & -0.0958 & 0.2968 & -0.0556 \\ -0.2135 & -0.1006 & -0.0082 & 0.121 & 0.1607 & 0.1414 & 0.0312 & -0.1667 & 0.0346 \\ 0.015 & 0.0061 & -0.001 & -0.0108 & -0.0136 & -0.0094 & 0.0006 & 0.0174 & -0.0045 \\ -0.0003 & -0.0001 & 0 & 0.0002 & 0.0003 & 0.0002 & -0.0001 & -0.0004 & 0.0002 \end{bmatrix},$$

Since

$$U_p \rightarrow \begin{bmatrix} -3.6 \cdot 10^{-5} & 0.005599 & 0.200277 & 0.709844 \\ -0.000286 & 0.020471 & 0.314476 & 0.437433 \\ -0.00096 & 0.043551 & 0.399688 & 0.210535 \\ -0.004425 & 0.109901 & 0.492289 & -0.119144 \\ -0.007637 & 0.150953 & 0.504246 & -0.228135 \\ -0.060915 & 0.437681 & 0.239075 & -0.295846 \\ -0.118905 & 0.557123 & -0.007317 & -0.092573 \\ -0.281683 & 0.617251 & -0.373393 & 0.307689 \\ -0.950119 & -0.282556 & 0.089436 & -0.058651 \end{bmatrix},$$

$$V_p \rightarrow \begin{bmatrix} -5.0 \cdot 10^{-5} & 0.008703 & 0.31918 & 0.947654 \\ -0.001292 & 0.100195 & 0.942642 & -0.318412 \\ -0.0353 & 0.994305 & -0.097704 & 0.023774 \\ -0.999376 & -0.035251 & 0.002217 & -0.000476 \end{bmatrix}$$

Condition number of the kernel:

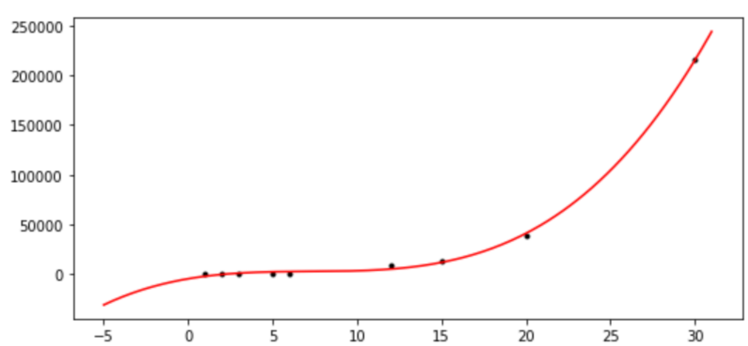
$$k_{\mathbf{frob}}(\mathbf{G}) = \|\mathbf{G}\|_F \|\mathbf{G}^{-g}\|_F = 31314.16.$$

Thus, we get the parameter estimates:

$$\mathbf{m}^{\mathbf{est}} = \mathbf{G}^{-\mathbf{g}}\mathbf{d} = \begin{bmatrix} -4526.12816842002 \\ 2881.04015282435 \\ -385.87700616485 \\ 17.7878675370993 \end{bmatrix},$$

$$\mathbf{d}^{\mathbf{pre}} = \mathbf{G}\mathbf{m}^{\mathbf{est}} = \begin{bmatrix} -2013.17715422342 \\ -165.252947133924 \\ 1124.37165807106 \\ 2455.63088371789 \\ 2710.71991460492 \\ 5217.49988184133 \\ 11901.2006745641 \\ 41046.8127189212 \\ 214888.194369626 \end{bmatrix},$$

$$\rightarrow \mathbf{d} = -\,4526.128 + 2881.04x + -385.877x^2 + 17.788x^3,$$



Data Analysis Matrix:

$$\mathbf{N} = \mathbf{G}\mathbf{G}^{-\mathbf{g}} = \left(U_p \Sigma_p V_p^T\right)\left(V_p \Sigma_p^{-1} U_p^T\right) = U_p U_p^T \Rightarrow$$

$$\mathbf{N} \approx \begin{bmatrix} 0.544 & 0.3736 & 0.2297 & 0.0146 & -0.0601 & -0.1597 & -0.0641 & 0.1471 & -0.0253 \\ 0.3736 & 0.2907 & 0.2187 & 0.1049 & 0.0619 & -0.0453 & -0.0314 & 0.0299 & -0.003 \\ 0.2297 & 0.2187 & 0.206 & 0.1765 & 0.1601 & 0.0524 & 0.002 & -0.0573 & 0.012 \\ 0.0146 & 0.1049 & 0.1765 & 0.2686 & 0.292 & 0.2013 & 0.0692 & -0.1514 & 0.0242 \\ -0.0601 & 0.0619 & 0.1601 & 0.292 & 0.3292 & 0.2546 & 0.1024 & -0.1631 & 0.0231 \\ -0.1597 & -0.0453 & 0.0524 & 0.2013 & 0.2546 & 0.34 & 0.2767 & 0.107 & -0.0271 \\ -0.0641 & -0.0314 & 0.002 & 0.0692 & 0.1024 & 0.2767 & 0.3331 & 0.3516 & -0.0397 \\ 0.1471 & 0.0299 & -0.0573 & -0.1514 & -0.1631 & 0.107 & 0.3516 & 0.6944 & 0.0418 \\ -0.0253 & -0.003 & 0.012 & 0.0242 & 0.0231 & -0.0271 & -0.0397 & 0.0418 & 0.994 \end{bmatrix}$$

Parameter Analysis Matrix:

$$\mathbf{R} = \mathbf{G}^{-\mathbf{g}}\mathbf{G} = \left(V_p \Sigma_p^{-1} U_p^T\right)\left(U_p \Sigma_p V_p^T\right) = V_p V_p^T \Rightarrow \mathbf{R} = id \in \mathbb{R}^{4,4}.$$

Note: The data analysis matrix depends on the measure process and not on the measurements.

The rows of the data analysis matrix describe how well specific data can be retrieved in relation to adjacent measurements.

i.e. for the 2$^{\text{nd}}$ row of the data analysis matrix, we get:

$$\mathbf{d_2^{pre}} = \sum_{j=1}^{9} \mathbf{N_{2j}}\ \mathbf{d_j^{obs}} = 8.59294 + 14.24243 + 23.39864 + \ \ldots\ -655.10978 = -165.252947\ .$$

This can obviously be generalized:

$$\mathbf{d_i^{pre}} = \sum_{j=1}^{9} \mathbf{N_{ij}}\ \mathbf{d_j^{obs}}, \quad i = 1, \ldots, 9.$$

Note: The parameter analysis matrix does not depend on the parameters, but only to an a-priori information (i.e mean value, confidence intervals, etc..). Since we don't have such an information here, it is expected to get the identity matrix.

$$\mathbf{spread(N)} = \|\mathbf{N} - \mathbf{I}\|_F^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}[\mathbf{N_{ij}} - \mathbf{I_{ij}}]^2 \quad \textbf{(Frobenius norm)}$$

$$= 5\ .$$

$$\left(\|\mathbf{N} - \mathbf{I}\|_2^2 = \rho\Big((\mathbf{N} - \mathbf{I})^{\mathbf{T}}(\mathbf{N} - \mathbf{I})\Big) = \sigma_{\max}^2(\mathbf{N} - \mathbf{I}) = 1\right)$$

We calculate the parameter variance matrix:

$$[\mathbf{cov_u m}] = \mathbf{G^{-g}} \underbrace{\mathbf{cov_u d}}_{\mathbf{id}} \left(\mathbf{G^{-g}}\right)^{\mathbf{T}} = \begin{bmatrix} 1.065704 & -0.348196 & 0.025741 & -0.000513 \\ -0.348196 & 0.146175 & -0.011671 & 0.000241 \\ 0.025741 & -0.011671 & 0.000977 & -2.1 \cdot 10^{-5} \\ -0.000513 & 0.000241 & -2.1 \cdot 10^{-5} & 0 \end{bmatrix},$$

Dirichlet spread functions:

$$\rightarrow \mathbf{size}\big([\mathbf{cov_u m}]\big) = \sum_{j=1}^{M} [\mathbf{cov_u m}]_{jj} = 1.065704 + 0.146175 + \cdots + 0 = 1.212856 \ .$$

$$[\mathbf{cov_u m^{est}}] = \mathbf{G^{-g}} \underbrace{\mathbf{cov_u d}}_{\mathbf{id}} \left(\mathbf{G^{-g}}\right)^{\mathbf{T}} + [\mathbf{I} - \mathbf{R}][\mathbf{cov_u m}][\mathbf{I} - \mathbf{R}]^{\mathbf{T}} =$$

$$= \begin{bmatrix} 1.065704 & -0.348196 & 0.025741 & -0.000513 \\ -0.348196 & 0.146175 & -0.011671 & 0.000241 \\ 0.025741 & -0.011671 & 0.000977 & -2.1 \cdot 10^{-5} \\ -0.000513 & 0.000241 & -2.1 \cdot 10^{-5} & 0 \end{bmatrix}.$$
