Numerical Linear Algebra Homework Project 2: Least Squares, Orthogonalization, and the SVD

Catalano Giuseppe, Cerrato Nunzia

1 Problem 1

(1) Suppose we are given m pairs of data points, $(x_1, y_1), \dots, (x_m, y_m)$. We want to find a linear combination of prescribed functions $\phi_1, \dots, \phi_n$ whose values at the points $x_i \in [a, b]$, $1 \le i \le m$, approximate the values $y_1, \dots, y_m$ as well as possible. More precisely, the problem is to find a function of the form $f(x) = \alpha_1\phi_1(x) + \dots + \alpha_n\phi_n(x)$ such that

$$\sum_{i=1}^{m} [y_i - f(x_i)]^2 \le \sum_{i=1}^{m} [y_i - g(x_i)]^2 \quad \forall g \in \mathrm{Span}(\phi_1, \dots, \phi_n), \tag{1}$$

where, usually, $m > n$. It is possible to rephrase the problem as:

$$f = \arg \min_{f \in \mathrm{Span}(\phi_1,\dots,\phi_n)} \sum_{i=1}^{m} [y_i - f(x_i)]^2 . \tag{2}$$

Now we can define a column vector $\mathbf{z} \in \mathbb{R}^n$ such that:

$$[\mathbf{z}]_i = \alpha_i \tag{3}$$

and a matrix A such that:

$$[A\mathbf{z}]_i = f(x_i) = \alpha_1\phi_1(x_i) + \dots + \alpha_n\phi_n(x_i). \tag{4}$$

In this way, the element of the $i - th$ row and $j - th$ column of the matrix A is:

$$[A]_{ij} = a_{ij} = \phi_j(x_i). \tag{5}$$

Finally, defining a column vector $\mathbf{b} \in \mathbb{R}^n$ such that:

$$[\mathbf{b}]_i = y_i \tag{6}$$

we can rewrite the (2) as follows:

$$\tilde{\mathbf{z}} = \arg \min_{\mathbf{z} \in \mathbb{R}^n} \|\mathbf{b} - A\mathbf{z}\|_2^2 = \arg \min_{\mathbf{z} \in \mathbb{R}^n} \|\mathbf{b} - A\mathbf{z}\|_2, \tag{7}$$

where the function f can be built from $\tilde{\mathbf{z}}$.
(2) Now we suppose to take $\phi_k = x^{k-1}$, $1 \le k \le n$. Under this assumption, the matrix A takes the form:

$$A = \begin{bmatrix} x_1^0 & \cdots & x_1^{n-1} \\ \vdots & \ddots & \vdots \\ x_m^0 & \cdots & x_m^{n-1} \end{bmatrix} . \tag{8}$$

We want to prove that, assuming that $x_i \ne x_j$ for $i \ne j$, A has full rank: $\mathrm{rank}(A) = n$.
Proof: Proving that $\mathrm{rank}(A) = n$ is equivalent to prove that $\dim(\ker(A)) = 0$, that means

that $\nexists \mathbf{v} \in \mathbb{R}^n$ s.t. $\mathbf{v} \in \ker(A)$. We want to prove this statement by contraddiction, therefore, we look for a vector $\mathbf{v} \in \mathbb{R}^n$, with $\mathbf{v} \neq \underline{0}$, such that $A\mathbf{v} = \underline{0}$, that means:

$$\begin{cases} v_1 x_1^0 + \cdots + v_n x_1^{n-1} = 0 \\ \vdots \\ v_1 x_m^0 + \cdots + v_n x_m^{n-1} = 0 \end{cases} . \tag{9}$$

Defining the polynomial

$$p_{\mathbf{v}}^{(n-1)}(x) = \sum_{i=1}^{n} v_i x^{i-1} \tag{10}$$

we can observe that, for any choice of $\mathbf{v} \neq \underline{0}$, $p_{\mathbf{v}}^{(n-1)}(x)$ admits at most $n-1$ different roots, therefore $\nexists \mathbf{v} \neq \underline{0}$ such that $A\mathbf{v} = \underline{0}$. This concludes the proof. $\square$

(3) Consider the problem of finding the best fit with a quadratic function $f(x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2$ for the following data:

x_i	8	10	12	16	20	30	40	60	100
y_i	0.88	1.22	1.64	2.72	3.96	7.66	11.96	21.56	43.16

In the following we report the code that solves the normal equations:

$$A^T A \mathbf{v} = A^T \mathbf{b} \tag{11}$$

using the Cholesky factorization algorithm and then compares the result with the one found using the QR factorization of the matrix A.
*** INSERIRE CODICE ***
(4)

Problem 3

(1) Let $A \in \mathbb{R}^{m \times n}$, with singular value decomposition $A = U\Sigma V^T = \sum_{i=1}^{n} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ and $\mathrm{rank}(A) = n$. Express the singular values and singular vectors of the following matrices in terms of those of A.

(a) $(A^T A)^{-1}$

$$(A^T A)^{-1} = (V^T)^{-1} \Sigma^{-2} V^{-1} = V \Sigma^{-2} V^{-1}$$

- Both left and right singular vectors of $(A^T A)^{-1}$ are equal to the right singular vectors of A;
- Singular values of $(A^T A)^{-1}$ are equal to singular values of A raised to the -2 power.

(b) $(A^T A)^{-1} A^T$

$$(A^T A)^{-1} A^T = V \Sigma^{-2} V^T V \Sigma U^T = V \Sigma^{-1} U^{-1}$$

- Left singular vectors of $(A^T A)^{-1} A^T$ are equal to the right singular vectors of A;
- Right singular vectors of $(A^T A)^{-1}$ are equal to the left singular vectors of A;
- Singular values of $(A^T A)^{-1}$ are equal to the inverse of singular values of A.

(c) $A(A^T A)^{-1}$

$$A(A^T A)^{-1} = U \Sigma V^T V \Sigma^{-2} V^{-T} = U \Sigma^{-1} V^T$$

- Left and right singular vectors of $(A^T A)^{-1}$ coincide with that of A;
- Singular values of $(A^T A)^{-1}$ are equal to the inverse of singular values of A.

(d) $A(A^T A)^{-1} A^T$

$$A(A^T A)^{-1} A^T = U\Sigma(V^T)V\Sigma^{-2}V^T V\Sigma U^T = \mathbb{1}$$

(2)

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \tag{12}$$

$$A^T A = \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} \tag{13}$$

$$\det(A^T A - \lambda \mathbb{1}) = 0 \tag{14}$$

$$\lambda_{1,2} = \frac{9 \pm \sqrt{65}}{2} \tag{15}$$

$$k_2(A) = \frac{\sigma_{\max}}{\sigma_{\min}} = \sqrt{\frac{9 + \sqrt{65}}{9 - \sqrt{65}}} \tag{16}$$