

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 ϕ_1, \dots, ϕ_n whose values at the points $x_i \in [a, b]$, $1 \leq i \leq m$, 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

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

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

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

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

$$[\mathbf{z}]_i = \alpha_i \quad (3)$$

and a matrix A such that:

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

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

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

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

$$[\mathbf{b}]_i = y_i \quad (6)$$

we can rewrite the (2) as follows:

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

where the function f can be built from $\tilde{\mathbf{z}}$.

(2) Now we suppose to take $\phi_k = x^{k-1}$, $1 \leq k \leq n$. Under this assumption, the matrix A takes the form:

$$A = \begin{bmatrix} x_1^0 & \dots & x_1^{n-1} \\ \vdots & \ddots & \vdots \\ x_m^0 & \dots & x_m^{n-1} \end{bmatrix}. \quad (8)$$

We want to prove that, assuming that $x_i \neq x_j$ for $i \neq j$, A has full rank: $\text{rank}(A) = n$.

Proof: Proving that $\text{rank}(A) = n$ 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:

$$\begin{cases} v_1x_1^0 + \dots + v_nx_1^{n-1} = 0 \\ \vdots \\ v_1x_m^0 + \dots + v_nx_m^{n-1} = 0 \end{cases} \quad (9)$$

Defining the polynomial

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

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)