1 Data fitting: Exact or LS Approximation

Provided with a finite (training) set of distinct (training) sample points in a dataset \( D \subset \mathbb{R}^2 \),
\[(x_i, y_i), \quad i = 0, 1, 2, \ldots, m, \quad x_i \in [a, b], \ y_i \in \mathbb{R}.
\] (1)

Let \( x = [x_1, \cdots, x_m]^T \). Let \( y = [y_1, \cdots, y_m]^T \). Denote by \( C[a, b] \) the vector space of continuous functions on \([a, b]\).

Model architecture I.

Let \( S[a,b] = \text{span}(f_j \mid j = 1 : n) \) be the subspace on the field \( \mathbb{R} \), where \( f_j \) are linearly independent functions in \( C[a, b] \), not necessarily polynomials. That is, \( S[a, b] \) is a chosen model space.

1. Describe the representation of the interpolant in the model space in terms of the spanning/coding functions \( \{f_j, j = 1 : n\} \).

2. (Y/N/M)
The fitting equations in matrix form can be expressed as follows,
\[A[\{x_i\}, \{f_j\}] \hat{c}_f = y,\] (2)
where
\[A(i, j) = f_j(x_i), \quad (i, j) \in \{1, \cdots, m\} \times \{1, \cdots, n\}\]

Claim: The matrix is of full column rank, i.e., if a solution exists, it is unique.

3. (Y/N/M)
Let \( \{q_j(x) \mid j = 1, \cdots, n\} \) be an orthogonal basis for \( S[a,b] \) obtained analytically by the Gram-Schmidt procedure. Then, the interpolation equations in matrix expression are as follows,
\[A[\{x_i\}, \{q_j\}] \hat{c}_q = y,\] (3)

Claim: The columns of matrix are mutually orthogonal with respect to the conventional inner product in \( \mathbb{R}^m \).

(Optional.) Describe the following relations in theory,
(a) between the two interpolation matrices \( A[\{x_i\}, \{f_j\}] \) and \( A[\{x_i\}, \{q_j\}] \)
(b) between the two coefficient vectors \( \hat{c}_f \) and \( \hat{c}_q \)
(c) between the two interpolants \( p_f(x) = \sum_j c_f(j)f_j(x) \) and \( p_q(x) = \sum_j c_q(j)q_j(x) \)
(d) between the matrix condition numbers (∗)

4. (Y/N/M)
Consider the LS-approximation counterparts of (2) and (3), respectively.
\[(\hat{c}_f^*, r_f^*) := \min \|A[\{x_i\}, \{f_j\}] \hat{c}_f - y\|^2\] (4)
\[(\hat{c}_q^*, r_q^*) := \|A[\{x_i\}, \{q_j\}] \hat{c}_q - y\|^2\] (5)
Here, the norm is induced by an inner product in $\mathbb{R}^m$ and common to both the LS problems, the outcome for each consists of a minima and its residual.

Claim: The two LS fitting residuals are of the same length, $||r_f^*|| = r_q^*$. The LS interpolants $p_f(x)$ and $p_q(x)$ are not necessarily the same.

**Model architecture II.**

Consider a fitting model with *domain decomposition*: $[a, b] = \bigcup_i \Delta_i$, where $\Delta_i$ are sub-intervals (sub-domains) with non-overlapping interior. The interpolant model has the following piecewise expression:

$$f(x) = \sum_{k=1}^{p} \chi_k(x) \cdot f_k(x), \quad x \in [a, b]$$

$$\Delta_k = [s_{k-1}, s_k], \quad a = s_0 < s_1 < \cdots < s_p = b, \quad (6)$$

$$f_k(x) = \sum_{j=1}^{n_k} c_{k,j} \cdot b_j(x), \quad x \in \Delta_k, \quad k = 1 : p,$$

where $\chi_k(x)$ is the indicator of $\Delta_k$, $n_k$ is the local degree of freedom (or regulation) for the interpolant piece $f_k(x)$ over $\Delta_k$. Typically, $n_k$ are small. Once $\{b_j\}$ and $\{n_k\}$ are prescribed, the interpolant $f$ is determined by the coefficients $\{c_{kj}, j = 1 : n_k, \quad k = 1 : p\}$. The set of subdivision points $\{s_j\}$ may or may not be a subset of $\{x_i\}$.

**Piecewise-polynomial interpolation**

Consider piece-wise polynomial interpolation.

The interpolation equations and smooth conditions in matrix expression for the piecewise polynomial interpolation at data $(x, y)$, can be described as follows,

$$\tilde{A} \mathbf{c} = \begin{pmatrix} A(\{x_i, i = 1 : m\} | \{b_j(\Delta_k)\}) \\ B_1(\{s_i, i = 1 : p-1\} | \{b_j(\Delta_k)\}) \\ B_2(\{a, b\} | \{b_j(\Delta_k), k = 1, p\}) \end{pmatrix} \mathbf{c} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \\ \mathbf{z} \end{pmatrix} \quad (7)$$

where $A$, for the fitting equations, has $m$ rows and $n = \sum_{k=1}^{p} n_k$ columns, $B_1$ is for smooth conditions at the sub-intervals boundary points $\{s_j\}$, and $B_2 \mathbf{c} = \mathbf{z}$ is for additional regulation conditions, typically at the end points, so that the interpolant solution is unique if exists.

5. (a) (Y/N/M) The interpolant, if exists, is guaranteed continuous when $\{s_j\} \subset \{x_i\}$.

(b) (Y/N/M) Matrix $A$ is sparse when $p$ is relatively large, each row has at most $\max(n_k)$ nonzero elements.

(c) (Y/N/M) If $B_1 \neq 0$, there are $(p - 1)$ rows for each of the smooth conditions in successively higher order, and each row of $B_1$ has at most $2 \max(n_k)$ nonzero elements.

6. Piecewise cubic polynomial interpolation ($n_k = 3$) such that the interpolant has continuous second derivative. Specify three different sets of basis polynomials $\{b_0(x), b_1(x), b_2(x), b_3(x)\}$ for all sub-intervals $\Delta_k, \quad k = 1 : p$. Describe the specific expressions of the smooth conditions (or matrix $B_1$) for each case.
7. Convert exact piecewise fitting (7) to piecewise LS fitting.

Choose the one(s), or create one, that make reasonable sense to you:

(a) \[ \text{arg min} \| \tilde{A} c - y \|^2 \]

(b) \[ \text{arg min} \| A c - y \|^2 \]

\[ \text{s.t.} \]

\[ B_1 c = 0 \]

\[ B_2 c = z \]

1.1 Matrix factorization

Let \( A \in \mathbb{R}^{m \times n} \). The LU and QR factorizations are commonly known and used:

\[
\begin{align*}
\text{LU:} & \quad AP = LU \\
\text{QR:} & \quad AP = QR
\end{align*}
\]

(10)

where \( P \) is a permutation matrix for reordering the columns. If the factorization is obtained numerically, the complexity in terms of arithmetic operations is \( O(mn \cdot \min(m, n)) \).

In general, \( A \) can be factored into sparser matrices as follows,

\[ A = A_1 \cdots A_k \]

(11)

where each factor matrix is orthogonal (square), diagonal (square), or triangular (lower, or upper, not necessarily square). Permutation matrices are special orthogonal matrices. In fact, the LU factors or the QR factors can be expressed by sparser factors. The DFT matrix of order \( n \) is factored into \( \log(n) \) sparse factors.

8. Describe a high-level procedure to determine whether or not there exists a solution to \( Ax = b \), using the factorization of (11), and get a solution if it exists.

9. Describe a high-level procedure to obtain an LS solution to \( \min_x \| Ax - b \|^2 \) using the QR factorization.
Analytical factorization: case study

Let $\mathcal{V}(\{x_i, i = 1 : n + 1\} \mid \{x^j, j = 0 : n\})$ be the Vandermonde matrix with monomial basis. The matrix is full. Let $\mathcal{N}(\{x_i, i = 1 : n + 1\} \mid \{b_j(x), j = 0 : n\})$ be the interpolation matrix with Newton’s basis polynomials. It is lower triangular.

10. Provide a brief reasoning: there is a unique nonsingular, upper triangular matrix $R$ for the basis transform,

$$[b_0(x), b_1(x), \cdots, b_n(x)] = [1, x, \cdots, x^n] R.$$  

(12)

Provide a brief reasoning for the following statement:

$$\mathcal{V} = \mathcal{N} R^{-1}$$  

(13)

is the LU factorization of the Vandermonde matrix.

Two immediate consequences:

(a) (Y/N/M) Matrix $\mathcal{V}$ is nonsingular iff $\mathcal{N}$ is nonsingular, and iff $\{x_i, i = 1 : n + 1\}$ are distinct.

(b) (Y/N/N) The analytical LU factorization of $\mathcal{V}$, in place of a numerical one, enables the reduction in complexity from $O(n^3)$ to $O(n^2)$ for solving the Vandermonde system $Ax = b$ with $A = \mathcal{V}$ or $A = \mathcal{V}^T$.

(Optional.)

Find a sparse factorization of $R$ or/and a sparse factorization of $\mathcal{N}$. The sparse factorizations lead to compressive representation and consequently a reduction in memory space from $O(n^2)$ to $O(n)$.