Numerical Methods

Approximation

Basic Concepts

2023-2024

[DΣIM]

Approximation

- A common problem, is to approximate a function f, by a member f^* of a class of functions easier to work with. For example, using polynomials, rational functions, or trigonometric polynomials.
- Each function in the class is specified by the numerical values of several parameters. We will restrict ourselves to functions of one variable defined in a closed interval.

Linear Approximation

- We shall be concerned with the problem of ***linear approximation***, i.e., a function f is to be approximated using a function f^* that can be expressed as a ***linear combination***:
- $$f^*(x) = f_{n+1}(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \cdots + a_n\varphi_n(x)$$
- of $n+1$ functions $\varphi_0, \varphi_1, \dots, \varphi_n$ chosen in advance within a certain set. The values $a_0, a_1, \dots, a_n$ are constants to be determined.

Linear Approximation

- If we take the functions $\varphi_i(x) = x^i$, the class of possible functions f^* will be the set of polynomials of degree n. The set $\{1, x, x^2, \dots, x^n\}$ is said to be a ***basis*** all polynomials of degree n.
- The function f can be given in different ways. A common situation is a table of values $f(x_0), f(x_1), \dots, f(x_m)$ defined in a set of distinct nodes $I = \{x_0, x_1, \dots, x_m\}$. In the second case, we know an analytic expression of the function f that needs to be approximated.

Normed Spaces

- We will need to work with linear spaces of functions to be able to build such linear combinations of the basis functions.
- We need also to be able to measure the distance between the function approximated f and the approximating function f^*.
- This measure is obtained using a metric giving a value to the difference $f - f^*$.

$$\left\| f - f^* \right\|$$

- These linear spaces must be also normed or inner product spaces.

Linear Approximation

- The general linear problem of best approximation is defined in a linear normed function space E and a linear subspace G in E. For any $f \in E$, we must consider the distance from f to G defined by some measure:

$$dist(f,G) = \inf_{g \in G} \|f - g\|$$

- We need to select the function $g \in G$ that gives the minimum distance between f and any element of G. If and element of g of G satisfies this property, we say that g is the ***best approximation*** of f. This approximation will depend on the ***norm*** chosen.

Norms and Seminorms

- The geometrical concept of the length of a vector has many natural applications within function spaces and approximation.
- Measuring the distance between two vectors $\boldsymbol{v}, \boldsymbol{w}$ is obtained by the length of the vector $\boldsymbol{v} - \boldsymbol{w}$. We would like to use the length or some distance to measure the goodness of an approximation.
- For this purpose, we will use different ***norms*** defined on the space of functions used.

Norms and Seminorms

- All ***norms*** are defined in a vector space $\boldsymbol{V}$ and satisfy the following properties:

$$a)\quad \|f\| \geq 0, \quad \forall f$$

$$b)\quad \|af\| = |a|\|f\|, \quad \forall a \in \mathbb{R}, \mathbb{C}$$

$$c)\quad \text{Triangle inequality: } \|f+g\| \leq \|f\| + \|g\|$$

$$d)\quad \|f\| = 0 \Leftrightarrow f = 0$$

- If condition d) is not satisfied, we have a ***seminorm***.

Discrete Norms

- The most common ***norms in a discrete space*** will be:
- A) Euclidean Norm:

$$\|f\|_2 = \left(\sum_{k=0}^{m} |f(x_k)|^2 \right)^{1/2}$$

- B) Weighted Euclidean Norm:

$$\|f\|_2 = \left(\sum_{k=0}^{m} w_k |f(x_k)|^2 \right)^{1/2}, \quad \sum_{k=0}^{m} w_k = 1$$

- The constants, w_k are the ***weights*** of the norm.
- C) Maximum Norm:

$$\|f\|_\infty = \max_{k=0,\ldots,m} |f(x_k)|$$

Continuous Norms

- In a closed interval, $[a, b]$ in the common norms are:
- A) Euclidean norm:

$$\|f\|_2 = \left(\int_a^b |f(x)|^2 dx \right)^{1/2}$$

- B) Weighted Euclidean norm:

$$\|f\|_2 = \left(\int_a^b w(x) |f(x)|^2 dx \right)^{1/2}$$

- Where the weight $w(x)$ is a continuous positive function
- C) Maximum Norm

$$\|f\|_\infty = \max_{x \in [a,b]} |f(x)|$$

Inner (Scalar) Product

- The concept of norm arises from the idea of ***inner product***. In a vector space $\boldsymbol{V}$ we can define the inner product as a function:

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

- With the properties:
- $\langle \boldsymbol{x}, \boldsymbol{y} + \lambda \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \lambda \langle \boldsymbol{x}, \boldsymbol{z} \rangle, \quad \forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{V}$
- $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{V}$
- $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0 \quad \forall \boldsymbol{x} \neq \boldsymbol{0}$

Inner (Scalar) Product

- The inner product and the norm are related by:

$$\boxed{\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}}$$

- The ***Cauchy-Schwarz relation*** derives from this definition. Given $x, y \in V$ and $\lambda \in \mathbb{R}$ we have

$$\|\mathbf{x} - \lambda \mathbf{y}\|^2 = \langle \mathbf{x} - \lambda \mathbf{y}, \mathbf{x} - \lambda \mathbf{y} \rangle = \|\mathbf{y}\|^2 - 2\lambda \langle \mathbf{x}, \mathbf{y} \rangle + \lambda^2 \|\mathbf{x}\|^2$$

- This quadratic equation in λ will have positive solution if the discriminant verifies:

$$4\langle \mathbf{x}, \mathbf{y} \rangle^2 - 4\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \leq 0$$

- Hence

$$\boxed{\left|\langle \mathbf{x}, \mathbf{y} \rangle\right| \leq \|\mathbf{x}\| \|\mathbf{y}\|}$$

Orthogonal Vectors

- Two vectors $x, y \in V$ are orthogonal if they verify:

$$\boxed{\langle \mathbf{x}, \mathbf{y} \rangle = 0}$$

- This definition generalize important results of Euclidean Geometry. If we set $\lambda = -1$ in the quadratic equation, we obtain the ***Pythagorean Theorem***:

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

- While combining the relations with $\lambda = -1$ and $\lambda = 1$ we obtain the ***parallelogram law***:

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{y} - \mathbf{x}\|^2 = 2\left(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\right)$$

Orthogonal Sets

- A set of vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n\} \in \boldsymbol{V}$ is an ***orthogonal set*** if we have $\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle = 0,\ \forall\, i \neq j$. Orthogonal sets are linear independent. If

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_n \mathbf{x}_n = 0$$

- Then
$$\langle \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_n \mathbf{x}_n, x_i \rangle = \lambda_i \left\| \mathbf{x}_i \right\|^2 = 0 \Leftrightarrow \lambda_i = 0$$

- Moreover, if we have $\|x_i\|^2 = \langle x_i, x_i \rangle = 1$ we obtain an ***orthonormal set***. If these sets expand ***V*** we have an ***orthogonal basis*** or and ***orthonormal basis***.

Existence

- The general problem of best approximation is ***well posed***. The solution exists and is unique under some conditions.
- ***Theorem:*** *If G is a finite-dimensional subspace in a normed linear space E, then each vector of E possesses at least one best approximation*. If $f \in E$, and $g \in G$, we have:

$$\|f-g\| \le \|f-0\| = \|f\|$$

- The set $K = \{g \in G : \|f-g\| \le f\}$. Is closed and bounded. As G is finite-dimensional, G is compact. This means that it will attain its infimum.

Strict Norms

- A norm is ***strict*** if the equality in the triangle inequality holds only if the two elements involved are linearly dependent. If $f, g \in V, f \neq 0, g \neq 0$ are such that:

$$\|f+g\| = \|f\| + \|g\|$$

- Then there exists a number $\lambda \in \mathbb{C}$ with $g = \lambda f$. In this case:

$$\|f+g\| = \|f+\lambda f\| = \|f\| + \|\lambda f\|$$

- But

$$\|f+\lambda f\| = |1+\lambda|\|f\|, \quad \|f\| + \|\lambda f\| = (1+|\lambda|)\|f\|$$

- And

$$|1+\lambda| = 1+|\lambda| \Leftrightarrow \lambda = |\lambda| \Rightarrow \lambda \in \mathbb{R}$$

Uniqueness

- The ***Euclidean norm is strict***. For strict norms, the best approximation f^*, is also ***unique***.
- If f_1^*, f_2^* are approximations with $\|f - f_1^*\| = \|f - f_2^*\| = \varepsilon$, then, it is not possible that:

$$\left\| f - \frac{f_1^* + f_2^*}{2} \right\| < \varepsilon$$

- And
- $$2\varepsilon = \left\| \left(f - f_1^* \right) + \left(f - f_2^* \right) \right\| = \left\| f - f_1^* \right\| + \left\| f - f_2^* \right\|$$
- $(f - f_1^*)$ and $(f - f_2^*)$ are independent with the same norm which implies that $f_1^* = f_2^*$

Approximation

- When we use a discrete set of nodes (x_k, y_k), $k = 0 \dots m$ we say that we are performing a ***discrete approximation***.
- We do not need to restrict ourselves to a discrete set of nodes. We could also approximate a function in the sense that

$$\left\| f - f^{*} \right\| = \left(\int_a^b (f(x) - f^{*}(x))^2 \, dx \right)^{1/2}$$

- Takes a minimum value. In this case we would be using a whole interval $[a, b]$ performing a ***continuous approximation***.

Discrete approximation

- In the case of ***discrete approximation,*** the function values can be grouped in a column vector:

$$f = \left(f(x_0), f(x_1), \ldots, f(x_n) \right)^T$$

- If we want to verify that $f^*(x_i) = f(x_i),\ i = 0,1,\ldots,m$ then we have the case of ***interpolation.***
- The interpolating polynomial is not difficult to compute, and it is possible to have an explicit way to bound the error involved.

Discrete Approximation

- In the case of interpolation, we need to compute the constants $a_0, a_1, \dots, a_n$ solving the linear system of equations:

$$\begin{aligned}
\varphi_0(x_0)a_0 + \varphi_1(x_0)a_1 + \cdots + \varphi_n(x_0)a_n &= f(x_0) \\
\varphi_0(x_1)a_0 + \varphi_1(x_1)a_1 + \cdots + \varphi_n(x_1)a_n &= f(x_1) \\
&\vdots \\
\varphi_0(x_m)a_0 + \varphi_1(x_m)a_1 + \cdots + \varphi_n(x_m)a_n &= f(x_m)
\end{aligned}$$

- Hence, $m = n$, while the solution is unique if the functions $\varphi_i(x)$ determine n independent vectors

$$\left(\varphi_i(x_0), \varphi_i(x_1), \ldots, \varphi_i(x_m)\right)^T \quad i = 0,1,\ldots,n$$

Discrete Approximation

- If $m > n$, only in exceptional cases we can get $f^*(x) = f(x)$ at all nodes. The system has more equations than unknowns. We say that the system is ***overdetermined***.
- In such a case, we can only ask for the equations to be satisfied approximately, minimizing the error

$$e_k = f(x_k) - f^*(x_k), \quad k = 0, \dots, m$$

- We obtain a ***smoothing*** of the data reducing the effect of random errors building a smooth curve through the data points.

Continuous approximation

- In the case of functions f given in analytic form, we will select a function f^* from a given function space that is sufficiently close to the function $f \approx f^*$
- This function space will be described by a ***set of basis functions*** $\varphi_0, \varphi_1, \dots, \varphi_n$ while the function f^* will be obtained by a ***linear combination*** of these functions. The constants $a_0, a_1, \dots, a_n$ will be selected using some proximity criterion that minimizes the error within the interval of definition of f:

$$e_k = f(x) - f^*(x), \quad x \in [a, b]$$

Geometric View

- In approximation methods we want to find a function

$$f^*(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \cdots + a_n\varphi_n(x)$$

- For which a given norm $\|f - f^*\|$ takes a value as small as possible. The set of basis functions $\varphi_0, \varphi_1, \dots, \varphi_n$ will be an independent set which spans a ***linear subspace***

$$S = \langle \varphi_0, \varphi_1, \dots, \varphi_n \rangle \subset \mathbb{F}$$

- Within a larger ***functional space***. We want to select the vector from this subspace that lies at the shortest distance from f.

Geometric View

- The solution to the approximation problem, when the Euclidean norm is used, is simply a generalization of the well-known geometrical fact from two and three dimensions: the shortest distance from a point to a linear subspace (plane, line) is the length of the ***vector which is perpendicular to the subspace***.
- The error vector $f - f^*$ ***will be perpendicular*** to the subspace generated by $\varphi_0, \varphi_1, \dots, \varphi_n$

Projections

- A projection is a linear map $\boldsymbol{P}: \boldsymbol{V} \rightarrow \boldsymbol{V}$ with the property that:

$$\mathbf{P}^2 = \mathbf{P}$$

- Given a vector $\boldsymbol{x} \in \boldsymbol{V}$ the projection will map this vector onto the subspace $\boldsymbol{U} = range(\boldsymbol{P})$ and leaves unchanged any vector that is already in this subspace. If $\boldsymbol{x} \in range(\boldsymbol{P})$ there is some vector $\boldsymbol{y} \in \boldsymbol{V}$ such that $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$. Then:

$$\mathbf{Px} = \mathbf{P}(\mathbf{Py}) = \mathbf{P}^2\mathbf{y} = \mathbf{Py} = \mathbf{x}$$

Projections

- If $\boldsymbol{V}$ is an inner product space, a projection satisfying the additional property $\boldsymbol{P} = \boldsymbol{P}^*$ is said to be an ***orthogonal projection***.
- Consider the difference between a general vector $\boldsymbol{y} \in \boldsymbol{V}$ and its projection $\boldsymbol{Py}$ onto $range(\boldsymbol{P})$. That is the residual vector defined as $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{Py}$. This vector is orthogonal to $range(\boldsymbol{P})$. If $\boldsymbol{x} \in range(\boldsymbol{P})$:

$$
\begin{aligned}
\langle \mathbf{x}, \mathbf{r} \rangle = \langle \mathbf{x}, \mathbf{y} - \mathbf{Py} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{Py} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{Px}, \mathbf{y} \rangle \\
= \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle = 0
\end{aligned}
$$

Projections

- ***Theorem**. Given an orthogonal projection* $\boldsymbol{P}$ *and an arbitrary vector* $\boldsymbol{y} \in \boldsymbol{V}$, *then the vector in* $range(\boldsymbol{P})$ *which is closest to* $\boldsymbol{y}$ *with respect to the Euclidean norm is given by* $\boldsymbol{Py}$.
- Let $y \in V$ and suppose that $x \in range(P)$, then:

$$
\begin{aligned}
\|\mathbf{x}-\mathbf{y}\|^2 - \|\mathbf{Py}-\mathbf{y}\|^2 &= \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle - \langle \mathbf{Py}-\mathbf{y}, \mathbf{Py}-\mathbf{y} \rangle \\
&= \langle \mathbf{x}, \mathbf{x} \rangle - 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle - \langle \mathbf{Py}, \mathbf{Py} \rangle + 2\langle \mathbf{y}, \mathbf{Py} \rangle - \langle \mathbf{y}, \mathbf{y} \rangle \\
&= \langle \mathbf{x}, \mathbf{x} \rangle - 2\langle \mathbf{Px}, \mathbf{y} \rangle + \langle \mathbf{Py}, \mathbf{Py} \rangle \\
&= \langle \mathbf{x}-\mathbf{Py}, \mathbf{x}-\mathbf{Py} \rangle = \|\mathbf{x}-\mathbf{Py}\|^2
\end{aligned}
$$

- So $\|\boldsymbol{x}-\boldsymbol{y}\| \geq \|\boldsymbol{Py}-\boldsymbol{y}\|$ with equality if and only if $\boldsymbol{x} = \boldsymbol{Py}$

Least Squares Method

- Let $I = \{x_0, x_1, \dots, x_m\}$ be the set of approximating nodes. And let $\{\varphi_j, j = 0, \dots n\}$ be the set of basis functions. We define F_n as the linear space defined by this set of basis functions:

$$F_n = \left\{ f_n \in F_n : f_n = \sum_{j=0}^{n} a_j \varphi_j, \;\; a_j \in \mathbb{R},\, j = 0, 1, \dots, n \right\}$$

- Given the function f to be approximated we want to search for a function $f_n \in F_n$ such that

$$\left\| f - f_n^* \right\|_2 = \min_{f_n \in F_n} \left\| f - f_n \right\|^2$$

Scalar Product

- Euclidean norms can be obtained from ***scalar products***. In the discrete case if f, g are functions defined on $I = \{x_0, x_1, \dots, x_m\}$ the scalar product is defined as:

$$\langle f, g \rangle = \sum_{k=0}^{m} f(x_k) g(x_k)$$

- And the Euclidean norm of a function f is defined as:

$$\|f\|_2^2 = \langle f, f \rangle$$

Scalar Product

- For the continuous case in the interval $[a, b]$ we have similar definitions. The scalar product is defined as:

$$\langle f, g \rangle = \int_a^b w(x) f(x) g(x) dx$$

- Where we have used a weight function $w(x)$ and the Euclidean norm of a function f is defined as:

$$\|f\|_{2,w}^2 = \langle f, f \rangle$$

Scalar Product

- Scalar products satisfy the following properties:
- Commutativity: $\langle f, g \rangle = \langle g, f \rangle$
- Linearity; $\langle c_1 f + c_2 g, h \rangle = c_1 \langle f, h \rangle + c_2 \langle g, h \rangle$
- Positivity: $\langle f, f \rangle \geq 0$
- From the rule of linearity if follows by induction that

$$\left\langle \sum_{j=0}^{n} c_j \varphi_j, \varphi_k \right\rangle = \sum_{j=0}^{n} c_j \left\langle \varphi_j, \varphi_k \right\rangle$$

Orthogonality

- Two functions f and g are said to be ***orthogonal*** if $\langle f, g\rangle = 0$. We represent the orthogonality by $f \perp g$. A finite or infinite sequence of functions $\varphi_0, \varphi_1, \dots, \varphi_n$ build an ***orthogonal system***, if $\langle \varphi_i, \varphi_j\rangle = 0$ for $i \neq j$ and $\|\varphi_i\| \neq 0$ for all $i = 0,1, \dots$ If in addition, $\|\varphi_i\| = 1$ for all $i = 0,1, \dots$, then the sequence is called an ***orthonormal system***. Then

$$\left\langle \varphi_i, \varphi_j \right\rangle = \delta_{ij}$$

- Where $\delta_{ij}$is the ***Kronecker delta***. Note that the common base of polynomials $\{1, x, x^2, \dots, x^n, \dots\}$ is ***not an orthogonal system***.

Least Squares

- The scalar product is a bilinear symmetric application. We have

$$
\begin{aligned}
\|f - f_n\|_2^2 &= \langle f - f_n, f - f_n \rangle \\
&= \langle f - f_n^* + f_n^* - f_n, f - f_n^* + f_n^* - f_n \rangle \\
&= \langle f - f_n^*, f - f_n^* \rangle + 2\langle f - f_n^*, f_n^* - f_n \rangle + \langle f_n^* - f_n, f_n^* - f_n \rangle \\
&= \|f - f_n^*\|_2^2 + 2\langle f - f_n^*, f_n^* \rangle - 2\langle f - f_n^*, f_n \rangle + \|f_n^* - f_n\|_2^2
\end{aligned}
$$

- If we choose $f_n^* \in F_n$ such that

$$
\langle f - f_n^*, f_n \rangle = 0 \quad \forall f_n \in F_n
$$

Least Squares

- As $f_n^* \in F_n$, we have:

$$\left\langle f - f_n^*, f_n \right\rangle = 0 = \left\langle f - f_n^*, f_n^* \right\rangle$$

- Then we obtain:

$$\left\| f - f_n \right\|_2^2 = \left\| f - f_n^* \right\|_2^2 + \left\| f_n^* - f_n \right\|_2^2$$

- And the minimum is reached when

$$\left\| f_n^* - f_n \right\|_2^2 = 0.$$

- Or $f_n^* = f_n$

Least Squares

- Our problem consists in finding $f_n^* \in F_n$ such that the ***error is orthogonal*** to any function within F_n

$$\left\langle f - f_n^*, f_n \right\rangle = 0 \qquad \forall f_n \in F_n$$

- This amounts to compute the constants $a_0^*, a_1^*, \dots, a_n^*$ that define the function f_n^* in terms of the basis functions $\varphi_0, \varphi_1, \dots, \varphi_n$

$$f_n^*(x) = \sum_{j=0}^{n} a_j^* \varphi_j(x)$$

- As F_n is the linear space generated by the set of basis functions $\varphi_0, \varphi_1, \dots, \varphi_n$ and considering that the scalar product is bilinear, we obtain:

$$\left\langle f - f_n^*, f_n \right\rangle = \left\langle f, f_n \right\rangle - \left\langle f_n^*, f_n \right\rangle$$

$$= \left\langle f, \sum_{i=0}^{n} a_i \varphi_i \right\rangle - \left\langle \sum_{j=0}^{n} a_j^* \varphi_j, \sum_{i=0}^{n} a_i \varphi_i \right\rangle = \sum_{i=0}^{n} a_i \left[\left\langle f, \varphi_i \right\rangle \right] - \sum_{i=0}^{n} a_i \left[\left\langle \sum_{j=0}^{n} a_j^* \varphi_j, \varphi_i \right\rangle \right]$$

$$= \sum_{i=0}^{n} a_i \left[\left\langle f, \varphi_i \right\rangle - \left\langle \sum_{j=0}^{n} a_j^* \varphi_j, \varphi_i \right\rangle \right] = \sum_{i=0}^{n} a_i \left[\left\langle f, \varphi_i \right\rangle - \sum_{j=0}^{n} a_j^* \left\langle \varphi_j, \varphi_i \right\rangle \right] = 0$$

Minimum Least Squares

- Then the condition for minimization is reduced to:

$$\sum_{j=0}^{n} a_j^* \left\langle \varphi_j, \varphi_i \right\rangle = \left\langle f, \varphi_i \right\rangle \qquad i = 0,1,\ldots,n$$

- This is a linear system of equations: $\mathbf{A} \cdot \mathbf{a}^* = \mathbf{b}$**,** where:

$$A = \left\langle \varphi_j, \varphi_i \right\rangle \qquad i, j = 0,1,\ldots,n$$

$$a^* = a_j^*, \qquad j = 0,1,\ldots,n$$

$$b = \left\langle f, \varphi_i \right\rangle, \qquad i = 0,1,\ldots,n$$

- From the properties of the scalar product, we know that **A** ***will be a symmetric positive definite matrix.***

Minimum Least Squares

- If we could use an ***orthogonal system***, however, we would obtain a diagonal system and the solution could be easily computed as:

$$a_j^* = \frac{b_i}{a_{jj}} = \frac{\langle f, \varphi_j \rangle}{\langle \varphi_j, \varphi_j \rangle}, \quad j = 0, 1, \ldots, n$$

- A family of basis functions verifying this property is an ***orthogonal basis***. These basis functions will give an optimal solution for the problem of minimum least squares.

Minimum Least Squares

- Another way to obtain the normal equations is to minimize the Euclidean difference between functions f and f^*. The distance between the two functions can be written as a function of the constants $a_0, a_1, \ldots, a_n$:

$$F\left(a_0, a_1, \ldots, a_n\right) = \sum_{k=0}^{m} \left[f(x_k) - \left(a_0 \varphi_0(x_k) + a_1 \varphi_1(x_k) + \ldots + a_n \varphi_n(x_k) \right) \right]^2$$

- The minimum at $a_0^*, a_1^*, \ldots, a_n^*$ will be reached when:

$$\frac{\partial F}{\partial a_i}\left(a_0^*, a_1^*, \ldots, a_n^*\right) = 0 \qquad i = 0, 1, \ldots, n$$

Minimum Least Squares

- This gives the equations:

$$-2\sum_{k=0}^{m}\left[f(x_k)-\left(a_0\varphi_0(x_k)+a_1\varphi_1(x_k)+\ldots+a_n\varphi_n(x_k)\right)\right]\varphi_i(x_k)=0,\quad i=0,1,\ldots,n$$

- Which can also be expressed as:

$$a_0^*\sum_{k=0}^{m}\varphi_0(x_k)\varphi_i(x_k)+a_1^*\sum_{k=0}^{m}\varphi_1(x_k)\varphi_i(x_k)+\cdots$$

$$\cdots+a_n^*\sum_{k=0}^{m}\varphi_n(x_k)\varphi_i(x_k)=\sum_{k=0}^{m}f(x_k)\varphi_i(x_k),\qquad i=0,1,\ldots,n$$

Minimum Least Squares

- If we write this relation in compact form:

$$a_0^* \langle \varphi_0, \varphi_i \rangle + a_1^* \langle \varphi_1, \varphi_i \rangle + \cdots + a_n^* \langle \varphi_n, \varphi_i \rangle = \langle f, \varphi_i \rangle, \quad i = 0, 1, \ldots, n$$

- We recover the normal equations.
- The problem is formally solved if we fix the set of basis functions. If we set the polynomial basis $\{1, x, x^2, \ldots, x^n\}$, however, the matrix of the system is normally ***bad conditioned***, and the rounding errors will propagate easily through the solution

Polynomial Approximation

- In this case we fix the set of polynomials $\{1, x, x^2, \dots, x^n\}$ as the functional basis. Consider the discrete case, where we have a table of the function f as $f(x_0), f(x_1), \dots, f(x_m)$ defined in a set of distinct nodes $I = \{x_0, x_1, \dots, x_m\}$
- The function f will be approximated for a polynomial of degree n of the form:

$$y(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

- The constants $a_0, a_1, \dots, a_n$ must be determined using the normal equations.

Polynomial Equations

- The set of normal equations in this case are:

$$a_0 m + a_1 \sum_{i=0}^{m} x_i + a_2 \sum_{i=0}^{m} x_i^2 + \cdots + a_n \sum_{i=0}^{m} x_i^n = \sum_{i=0}^{m} f(x_i)$$

$$a_0 \sum_{i=0}^{m} x_i + a_1 \sum_{i=0}^{m} x_i^2 + a_2 \sum_{i=0}^{m} x_i^3 + \cdots + a_n \sum_{i=0}^{m} x_i^{n+1} = \sum_{i=0}^{m} x_i f(x_i)$$

$$a_0 \sum_{i=0}^{m} x_i^2 + a_1 \sum_{i=0}^{m} x_i^3 + a_2 \sum_{i=0}^{m} x_i^4 + \cdots + a_n \sum_{i=0}^{m} x_i^{n+2} = \sum_{i=0}^{m} x_i^2 f(x_i)$$

$$\vdots$$

$$a_0 \sum_{i=0}^{m} x_i^n + a_1 \sum_{i=0}^{m} x_i^{n+1} + a_2 \sum_{i=0}^{m} x_i^{n+2} + \cdots + a_n \sum_{i=0}^{m} x_i^{2n} = \sum_{i=0}^{m} x_i^n f(x_i)$$

Polynomial Approximation

- An if we write these equations in matrix form as $\boldsymbol{Ba} = \boldsymbol{y}$, we obtain:

$$\begin{pmatrix} m & \sum_{i=0}^{m} x_i & \sum_{i=0}^{m} x_i^2 & \cdots & \sum_{i=0}^{m} x_i^n \\ \sum_{i=0}^{m} x_i & \sum_{i=0}^{m} x_i^2 & \sum_{i=0}^{m} x_i^3 & \cdots & \sum_{i=0}^{m} x_i^{n+1} \\ \sum_{i=0}^{m} x_i^2 & \sum_{i=0}^{m} x_i^3 & \vdots & \cdots & \sum_{i=0}^{m} x_i^{n+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{i=0}^{m} x_i^n & \sum_{i=0}^{m} x_i^{n+1} & \sum_{i=0}^{m} x_i^{n+2} & \cdots & \sum_{i=0}^{m} x_i^{2n} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{m} f(x_i) \\ \sum_{i=0}^{m} x_i f(x_i) \\ \sum_{i=0}^{m} x_i^2 f(x_i) \\ \vdots \\ \sum_{i=0}^{m} x_i^n f(x_i) \end{pmatrix}$$

- The resulting system matrix is ***highly ill-conditioned***, this limits the approximation to polynomials up to degree 3-4.

Polynomial Approximation

- The matrix of the system of normal equations, ***B***, in the least squares polynomial approximation is related to another matrix, called the ***design matrix:***

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_0 & x_1 & x_2 & \cdots & x_m \\ x_0^2 & x_1^2 & x_2^2 & \cdots & x_m^2 \\ \vdots & \vdots & & & \vdots \\ x_0^n & x_1^n & x_2^n & \ldots & x_m^n \end{pmatrix}$$

- We have $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}$

Polynomial Approximation

- Moreover, we have:

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_0 & x_1 & x_2 & \cdots & x_m \\ x_0^2 & x_1^2 & x_2^2 & \cdots & x_m^2 \\ \vdots & \vdots & & & \vdots \\ x_0^n & x_1^n & x_2^n & \dots & x_m^n \end{pmatrix} \begin{pmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \\ \vdots \\ f(x_m) \end{pmatrix} = \begin{pmatrix} \sum_{i=0}^{m} f(x_i) \\ \sum_{i=0}^{m} x_i f(x_i) \\ \sum_{i=0}^{m} x_i^2 f(x_i) \\ \vdots \\ \sum_{i=0}^{m} x_i^n f(x_i) \end{pmatrix}$$

- And we can write:

$$\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{a} = \mathbf{B}\mathbf{a} = \mathbf{A}\mathbf{y}$$

Singular Values

- The matrix $\mathbf{B} = \mathbf{A}\mathbf{A}^{\mathbf{T}}$ is positive definite. This class of matrices can be diagonalized by an orthogonal matrix $\mathbf{P}$ such that:

$$\mathbf{P}\mathbf{B}\mathbf{P}^{\mathrm{T}} = \mathbf{P}\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}} = \mathbf{D}$$

- With $\mathbf{P}\mathbf{P}^{\mathbf{T}} = \mathbf{I}$.
- All the eigenvalues of $\mathbf{B}$ will be nonnegative. This means that we can define a matrix $\mathbf{S} = \sqrt{\mathbf{D}}$ or $\mathbf{S}^2 = \mathbf{D}$. The diagonal elements of $\mathbf{S}$ are the ***singular values*** of $\mathbf{A}$.

Singular Values

- Then the ill-conditioned system defined by:

$$\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{a} = \mathbf{B}\mathbf{a} = \mathbf{A}\mathbf{y}$$

- Can be solved as:

$$\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{a} = \mathbf{P}^{\mathrm{T}}\mathbf{D}\mathbf{P}\mathbf{a} = (\mathbf{S}\mathbf{P})^{\mathrm{T}}(\mathbf{S}\mathbf{P})\mathbf{a} = \mathbf{A}\mathbf{y}$$

$$\mathbf{a} = \mathbf{P}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{P}\mathbf{A}\mathbf{y}$$

- This technique is known as ***singular value decomposition***.