Numerical Analysis
The method of least squares

by Csaba Gáspár

Széchenyi István University

2020, autumn semester

The method of least squares

Let $A \in \mathbf{M}_{N\times N}$ be a **regular** matrix. Let $b \in \mathbf{R}^N$ be a vector and consider the linear system of equations:

$$Ax = b$$

Denote by x^* the exact solution.

The method of least squares

The exact solution x^* minimizes the square of the norm of the residual vector, i.e. it minimizes the following functional:

$$F(x) := ||Ax - b||^2 = ||A(x - x^*)||^2 = ||x - x^*||^2_{A^*A}$$

The minimizing vector equals to that of the energetic functional of the **Gaussian normal equation**:

$$A^*Ax = A^*b$$

The method of least squares

Let $A \in \mathbf{M}_{N\times N}$ be a **regular** matrix. Let $b \in \mathbf{R}^N$ be a vector and consider the linear system of equations:

$$Ax = b$$

Denote by x^* the exact solution.

The method of least squares

The exact solution x^* minimizes the square of the norm of the residual vector, i.e. it minimizes the following functional:

$$F(x) := ||Ax - b||^2 = ||A(x - x^*)||^2 = ||x - x^*||^2_{A^*A}$$

The minimizing vector equals to that of the energetic functional of the **Gaussian normal equation**:

$$A^*Ax = A^*b$$

The method of least squares

Let $A \in \mathbf{M}_{N \times N}$ be a **regular** matrix. Let $b \in \mathbf{R}^N$ be a vector and consider the linear system of equations:

$$Ax = b$$

Denote by x^* the exact solution.

The method of least squares

The exact solution x^* minimizes the square of the norm of the residual vector, i.e. it minimizes the following functional:

$$F(x) := ||Ax - b||^2 = ||A(x - x^*)||^2 = ||x - x^*||^2_{A^*A}$$

The minimizing vector equals to that of the energetic functional of the **Gaussian normal equation**:

$$A^*Ax = A^*b$$

The method of least squares

Let $A \in \mathbf{M}_{N \times N}$ be a **regular** matrix. Let $b \in \mathbf{R}^N$ be a vector and consider the linear system of equations:

$$Ax = b$$

Denote by x^* the exact solution.

The method of least squares

The exact solution x^* minimizes the square of the norm of the residual vector, i.e. it minimizes the following functional:

$$F(x) := ||Ax - b||^2 = ||A(x - x^*)||^2 = ||x - x^*||^2_{A^*A}$$

The minimizing vector equals to that of the energetic functional of the **Gaussian normal equation**:

$$A^*Ax = A^*b$$

The method of least squares

Let $A \in \mathbf{M}_{N\times N}$ be a **regular** matrix. Let $b \in \mathbf{R}^N$ be a vector and consider the linear system of equations:

$$Ax = b$$

Denote by x^* the exact solution.

The method of least squares

The exact solution x^* minimizes the square of the norm of the residual vector, i.e. it minimizes the following functional:

$$F(x) := ||Ax - b||^2 = ||A(x - x^*)||^2 = ||x - x^*||^2_{A^*A}$$

The minimizing vector equals to that of the energetic functional of the **Gaussian normal equation**:

$$A^*Ax = A^*b$$

The method of least squares

Gradient method to the Gaussian normal equations

In each step: $\boxed{r^{(n)} := A^*(Ax^{(n)} - b)}$, and:

$$\boxed{x^{(n+1)} := x^{(n)} - \frac{||r^{(n)}||^2}{||Ar^{(n)}||^2} \cdot r^{(n)} \quad (n = 0, 1, 2, ...)}$$

The drawback of the use of the Gaussian normal equations is the fact that the symmetrized matrix A^*A is generally much more ill-conditioned than the original matrix A.

The method of least squares

Gradient method to the Gaussian normal equations

In each step: $\boxed{r^{(n)} := A^*(Ax^{(n)} - b)}$, and:

$$\boxed{x^{(n+1)} := x^{(n)} - \frac{||r^{(n)}||^2}{||Ar^{(n)}||^2} \cdot r^{(n)} \qquad (n = 0, 1, 2, ...)}$$

The drawback of the use of the Gaussian normal equations is the fact that the symmetrized matrix A^*A is generally much more ill-conditioned than the original matrix A.

The method of least squares

The least squares approach can be strongly generalized.
The matrix A need not be square (overdetermined, underdetermined systems)

Let $A \in \mathbf{M}_{N \times M}$ be a not necessarily square matrix. The solution of the system $Ax = b$ in the sense of least squares (or: the **generalized solution**) is called the vector which minimizes the norm of the residual

$$F(x) := ||Ax - b||^2$$

The generalized solution satisfies the Gaussian normal equations:

$$A^* Ax = A^* b$$

The method of least squares

The least squares approach can be strongly generalized. The matrix A need not be square (overdetermined, underdetermined systems)

Let $A \in \mathbf{M}_{N \times M}$ be a not necessarily square matrix. The solution of the system $Ax = b$ in the sense of least squares (or: the **generalized solution**) is called the vector which minimizes the norm of the residual

$$F(x) := ||Ax - b||^2$$

The generalized solution satisfies the Gaussian normal equations:

$$A^* A x = A^* b$$

The method of least squares

The least squares approach can be strongly generalized. The matrix A need not be square (overdetermined, underdetermined systems)

Let $A \in \mathbf{M}_{N \times M}$ be a not necessarily square matrix. The solution of the system $Ax = b$ in the sense of least squares (or: the **generalized solution**) is called the vector which minimizes the norm of the residual

$$F(x) := ||Ax - b||^2$$

The generalized solution satisfies the Gaussian normal equations:

$$A^* Ax = A^* b$$

The method of least squares

The least squares approach can be strongly generalized. The matrix A need not be square (overdetermined, underdetermined systems)

Let $A \in \mathbf{M}_{N\times M}$ be a not necessarily square matrix. The solution of the system $Ax = b$ in the sense of least squares (or: the **generalized solution**) is called the vector which minimizes the norm of the residual

$$F(x) := ||Ax - b||^2$$

The generalized solution satisfies the Gaussian normal equations:

$$A^*Ax = A^*b$$

The method of least squares. Linear regression

Let $x_1, x_2, ..., x_M$ be given numbers and let $y_1, y_2, ..., y_M$ be some associated values. Find a polynomial $p(x) = a_0 + a_1 x$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows: $a_0 + a_1 x_k = y_k$ for all $(k = 1, 2, ..., M)$.

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ ... & ... \\ 1 & x_M \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_M \end{pmatrix}$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min !$$

The method of least squares. Linear regression

Let $x_1, x_2, ..., x_M$ be given numbers and let $y_1, y_2, ..., y_M$ be some associated values. Find a polynomial $p(x) = a_0 + a_1 x$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows: $a_0 + a_1 x_k = y_k$ for all $(k = 1, 2, ..., M)$.

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_M \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_M \end{pmatrix}$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

The method of least squares. Linear regression

Let $x_1, x_2, ..., x_M$ be given numbers and let $y_1, y_2, ..., y_M$ be some associated values. Find a polynomial $p(x) = a_0 + a_1 x$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows: $a_0 + a_1 x_k = y_k$ for all $(k = 1, 2, ..., M)$.

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ ... & ... \\ 1 & x_M \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_M \end{pmatrix}$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

The method of least squares. Linear regression

Let $x_1, x_2, ..., x_M$ be given numbers and let $y_1, y_2, ..., y_M$ be some associated values. Find a polynomial $p(x) = a_0 + a_1 x$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows: $a_0 + a_1 x_k = y_k$ for all $(k = 1, 2, ..., M)$.

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ ... & ... \\ 1 & x_M \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_M \end{pmatrix}$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

The method of least squares. Linear regression

Let $x_1, x_2, ..., x_M$ be given numbers and let $y_1, y_2, ..., y_M$ be some associated values. Find a polynomial $p(x) = a_0 + a_1 x$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows: $a_0 + a_1 x_k = y_k$ for all $(k = 1, 2, ..., M)$.

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ ... & ... \\ 1 & x_M \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_M \end{pmatrix}$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

The method of least squares. Linear regression

$$F(a_0, a_1) := \sum_{k=1}^{M} (a_0 + a_1 x_k - y_k)^2 \to \min !$$

At the location of the minimum, the partial derivatives vanish:

$$\frac{\partial F}{\partial a_0} = \sum_{k=1}^{M} 2 \cdot (a_0 + a_1 x_k - y_k) \cdot 1 = 0$$

$$\frac{\partial F}{\partial a_1} = \sum_{k=1}^{M} 2 \cdot (a_0 + a_1 x_k - y_k) \cdot x_k = 0$$

whence a_0, a_1 can be computed.

The method of least squares. Linear regression

$$F(a_0, a_1) := \sum_{k=1}^{M} (a_0 + a_1 x_k - y_k)^2 \to \min !$$

At the location of the minimum, the partial derivatives vanish:

$$\frac{\partial F}{\partial a_0} = \sum_{k=1}^{M} 2 \cdot (a_0 + a_1 x_k - y_k) \cdot 1 = 0$$

$$\frac{\partial F}{\partial a_1} = \sum_{k=1}^{M} 2 \cdot (a_0 + a_1 x_k - y_k) \cdot x_k = 0$$

whence a_0, a_1 can be computed.

The method of least squares. Linear regression

The coefficients of the linear regression:

$$a_0 \cdot \sum_{k=1}^{M} 1 + a_1 \cdot \sum_{k=1}^{M} x_k = \sum_{k=1}^{M} y_k$$

$$a_0 \cdot \sum_{k=1}^{M} x_k + a_1 \cdot \sum_{k=1}^{M} x_k^2 = \sum_{k=1}^{M} x_k y_k$$

Quadratic regression

Let x_1, x_2, ..., x_M be given numbers and let y_1, y_2, ..., y_M be some associated values $(M \geq 3)$. Find a polynomial $p(x) = a_0 + a_1 x + a_2 x^2$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows:

$$a_0 + a_1 x_1 + a_2 x_1^2 = y_1$$

$$a_0 + a_1 x_2 + a_2 x_2^2 = y_2$$

$$\ldots$$

$$a_0 + a_1 x_M + a_2 x_M^2 = y_M$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

Quadratic regression

Let x_1, x_2, ..., x_M be given numbers and let y_1, y_2, ..., y_M be some associated values $(M \geq 3)$. Find a polynomial $p(x) = a_0 + a_1 x + a_2 x^2$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows:

$$a_0 + a_1 x_1 + a_2 x_1^2 = y_1$$
$$a_0 + a_1 x_2 + a_2 x_2^2 = y_2$$
$$\ldots$$
$$a_0 + a_1 x_M + a_2 x_M^2 = y_M$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

Quadratic regression

Let $x_1, x_2, \ldots, x_M$ be given numbers and let $y_1, y_2, \ldots, y_M$ be some associated values ($M \geq 3$). Find a polynomial $p(x) = a_0 + a_1 x + a_2 x^2$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows:

$$a_0 + a_1 x_1 + a_2 x_1^2 = y_1$$
$$a_0 + a_1 x_2 + a_2 x_2^2 = y_2$$
$$\ldots$$
$$a_0 + a_1 x_M + a_2 x_M^2 = y_M$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

Quadratic regression

Let x_1, x_2, ..., x_M be given numbers and let y_1, y_2, ..., y_M be some associated values ($M \geq 3$). Find a polynomial $p(x) = a_0 + a_1 x + a_2 x^2$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows:

$$a_0 + a_1 x_1 + a_2 x_1^2 = y_1$$
$$a_0 + a_1 x_2 + a_2 x_2^2 = y_2$$
$$\ldots$$
$$a_0 + a_1 x_M + a_2 x_M^2 = y_M$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

Quadratic regression

$$F(a_0, a_1, a_2) = ||A\underline{a} - \underline{y}||^2 = \sum_{k=1}^{M} (a_0 + a_1 x_k + a_2 x_k^2 - y_k)^2 \to \min!$$

$$\frac{\partial F}{\partial a_0} = \sum_{k=1}^{M} 2 \cdot (a_0 + a_1 x_k + a_2 x_k^2 - y_k) \cdot 1 = 0$$

$$\frac{\partial F}{\partial a_1} = \sum_{k=1}^{M} 2 \cdot (a_0 + a_1 x_k + a_2 x_k^2 - y_k) \cdot x_k = 0$$

$$\frac{\partial F}{\partial a_2} = \sum_{k=1}^{M} 2 \cdot (a_0 + a_1 x_k + a_2 x_k^2 - y_k) \cdot x_k^2 = 0$$

whence a_0, a_1, a_2 can be computed.

Quadratic regression

$$F(a_0,a_1,a_2) = ||A\underline{a}-\underline{y}||^2 = \sum_{k=1}^{M}(a_0+a_1x_k+a_2x_k^2-y_k)^2 \to \min!$$

$$\frac{\partial F}{\partial a_0} = \sum_{k=1}^{M} 2\cdot(a_0+a_1x_k+a_2x_k^2-y_k)\cdot 1 = 0$$

$$\frac{\partial F}{\partial a_1} = \sum_{k=1}^{M} 2\cdot(a_0+a_1x_k+a_2x_k^2-y_k)\cdot x_k = 0$$

$$\frac{\partial F}{\partial a_2} = \sum_{k=1}^{M} 2\cdot(a_0+a_1x_k+a_2x_k^2-y_k)\cdot x_k^2 = 0$$

whence a_0, a_1, a_2 can be computed.

Quadratic regression

The coefficients of the quadratic regression:

$$a_0 \cdot \left(\sum_{k=1}^{M} 1 \right) + a_1 \cdot \left(\sum_{k=1}^{M} x_k \right) + a_2 \cdot \left(\sum_{k=1}^{M} x_k^2 \right) = \sum_{k=1}^{M} y_k$$

$$a_0 \cdot \left(\sum_{k=1}^{M} x_k \right) + a_1 \cdot \left(\sum_{k=1}^{M} x_k^2 \right) + a_2 \cdot \left(\sum_{k=1}^{M} x_k^3 \right) = \sum_{k=1}^{M} x_k y_k$$

$$a_0 \cdot \left(\sum_{k=1}^{M} x_k^2 \right) + a_1 \cdot \left(\sum_{k=1}^{M} x_k^3 \right) + a_2 \cdot \left(\sum_{k=1}^{M} x_k^4 \right) = \sum_{k=1}^{M} x_k^2 y_k$$

Polynomial regression

Let x_1, x_2, ..., x_M be given numbers and let y_1, y_2, ..., y_M be some associated values $(M \geq p+1)$. Find a polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_p x^p$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows:

$$a_0 + a_1 x_1 + a_2 x_1^2 + ... + x_1^p = y_1$$
$$a_0 + a_1 x_2 + a_2 x_2^2 + ... + x_2^p = y_2$$
$$...$$
$$a_0 + a_1 x_M + a_2 x_M^2 + ... + x_M^p = y_M$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

Polynomial regression

Let x_1, x_2, ..., x_M be given numbers and let y_1, y_2, ..., y_M be some associated values $(M \geq p+1)$. Find a polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_p x^p$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows:

$$a_0 + a_1 x_1 + a_2 x_1^2 + ... + x_1^p = y_1$$

$$a_0 + a_1 x_2 + a_2 x_2^2 + ... + x_2^p = y_2$$

$$...$$

$$a_0 + a_1 x_M + a_2 x_M^2 + ... + x_M^p = y_M$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

Polynomial regression

Let x_1, x_2, ..., x_M be given numbers and let y_1, y_2, ..., y_M be some associated values $(M \geq p+1)$. Find a polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_p x^p$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows:

$$a_0 + a_1 x_1 + a_2 x_1^2 + ... + x_1^p = y_1$$

$$a_0 + a_1 x_2 + a_2 x_2^2 + ... + x_2^p = y_2$$

$$...$$

$$a_0 + a_1 x_M + a_2 x_M^2 + ... + x_M^p = y_M$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min !$$

Polynomial regression

Let $x_1, x_2, \dots, x_M$ be given numbers and let $y_1, y_2, \dots, y_M$ be some associated values $(M \geq p+1)$. Find a polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_p x^p$ such that it fits to the given data 'as exactly as possible'.

The exact fitting would be as follows:

$$a_0 + a_1 x_1 + a_2 x_1^2 + \dots + x_1^p = y_1$$
$$a_0 + a_1 x_2 + a_2 x_2^2 + \dots + x_2^p = y_2$$
$$\dots$$
$$a_0 + a_1 x_M + a_2 x_M^2 + \dots + x_M^p = y_M$$

There is no exact solution in general. It is worth defining the 'best fitting' by minimizing the norm of the residual:

$$F(\mathbf{a}) := ||A\mathbf{a} - \mathbf{y}||^2 \to \min!$$

Polynomial regression

At the minimum, the partial derivatives with respect to all a_m vanish:

$$\frac{\partial F}{\partial a_m} = \sum_{k=1}^{M} 2 \cdot \left(\sum_{j=0}^{p} a_j x_k^j - y_k \right) \cdot x_k^m = 0$$

The coefficients a_0, a_1,... , a_p satisfy the following system of equations:

$$\sum_{j=0}^{p} a_j \cdot \left(\sum_{k=1}^{M} x_k^{j+m} \right) = \sum_{k=1}^{M} x_k^m y_k \qquad (m = 0, 1, ..., p)$$

Overdetermined linear systems

Let $A \in \mathbf{M}_{N\times M}$ be (a necessarily square) matrix with $N > M$ (or even $N \gg M$). Let $b \in \mathbf{R}^N$ be a vector and consider the *overdetermined system of equations*:

$$Ax = b$$

No solution in general.

The vector $x^* \in \mathbf{R}^M$ is called a least squares solution or *generalized solution* if it minimizes the square of the Euclidean norm of the residual vector: $F(x) := ||Ax - b||^2$

Theorem: The generalized solution $x^* \in \mathbf{R}^M$ satisfies the *Gaussian normal equation*:

$$A^*Ax = A^*b$$

Overdetermined linear systems

Let $A \in \mathbf{M}_{N\times M}$ be (a necessarily square) matrix with $N > M$ (or even $N \gg M$). Let $b \in \mathbf{R}^N$ be a vector and consider the *overdetermined system of equations*:

$$Ax = b$$

No solution in general.

The vector $x^* \in \mathbf{R}^M$ is called a least squares solution or *generalized solution* if it minimizes the square of the Euclidean norm of the residual vector: $F(x) := ||Ax - b||^2$

Theorem: The generalized solution $x^* \in \mathbf{R}^M$ satisfies the *Gaussian normal equation*:

$$A^*Ax = A^*b$$

Overdetermined linear systems

Let $A \in \mathbf{M}_{N \times M}$ be (a necessarily square) matrix with $N > M$ (or even $N \gg M$). Let $b \in \mathbf{R}^N$ be a vector and consider the *overdetermined system of equations*:

$$Ax = b$$

No solution in general.

The vector $x^* \in \mathbf{R}^M$ is called a least squares solution or *generalized solution* if it minimizes the square of the Euclidean norm of the residual vector: $F(x) := ||Ax - b||^2$

Theorem: The generalized solution $x^* \in \mathbf{R}^M$ satisfies the *Gaussian normal equation*:

$$A^* Ax = A^* b$$

Overdetermined linear systems

Let $A \in \mathbf{M}_{N \times M}$ be (a necessarily square) matrix with $N > M$ (or even $N \gg M$). Let $b \in \mathbf{R}^N$ be a vector and consider the *overdetermined system of equations*:

$$Ax = b$$

No solution in general.

The vector $x^* \in \mathbf{R}^M$ is called a least squares solution or *generalized solution* if it minimizes the square of the Euclidean norm of the residual vector: $F(x) := ||Ax - b||^2$

Theorem: The generalized solution $x^* \in \mathbf{R}^M$ satisfies the *Gaussian normal equation*:

$$A^* Ax = A^* b$$

Overdetermined linear systems

Proof:

$$F(x) = ||Ax - b||^2 = \langle A^* Ax, x\rangle - 2\langle x, A^* b\rangle + ||b||^2$$

The gradient vector of F at x is: $\operatorname{grad} F(x) = 2A^* Ax - 2A^* b$

At the location of minimum, the gradient vector vanishes.

$A^* A \in \mathbf{M}_{M\times M}$, therefore the size of the matrix of the Gaussian normal equation is much less than that of the original system, if $N \gg M$.

The Gaussian normal equation may be severely ill-conditioned.

The method of least squares remains applicable when $N < M$ (underdetermined systems). Now the Gaussian normal equation (and possibly also the origial system) may have several solutions. Sometimes it is worth finding the solution with minimal Euclean norm.

Overdetermined linear systems

Proof:

$$F(x) = ||Ax - b||^2 = \langle A^*Ax, x\rangle - 2\langle x, A^*b\rangle + ||b||^2$$

The gradient vector of F at x is: $\operatorname{grad} F(x) = 2A^*Ax - 2A^*b$
At the location of minimum, the gradient vector vanishes.

$A^*A \in \mathbf{M}_{M\times M}$, therefore the size of the matrix of the Gaussian normal equation is much less than that of the original system, if $N \gg M$.

The Gaussian normal equation may be severely ill-conditioned.

The method of least squares remains applicable when $N < M$ (underdetermined systems). Now the Gaussian normal equation (and possibly also the origial system) may have several solutions. Sometimes it is worth finding the solution with minimal Euclean norm.

Overdetermined linear systems

Proof:

$$F(x) = ||Ax - b||^2 = \langle A^*Ax, x\rangle - 2\langle x, A^*b\rangle + ||b||^2$$

The gradient vector of F at x is: $\operatorname{grad} F(x) = 2A^*Ax - 2A^*b$
At the location of minimum, the gradient vector vanishes.

$A^*A \in \mathbf{M}_{M\times M}$, therefore the size of the matrix of the Gaussian normal equation is much less than that of the original system, if $N \gg M$.

The Gaussian normal equation may be severely ill-conditioned.

The method of least squares remains applicable when $N < M$ (underdetermined systems). Now the Gaussian normal equation (and possibly also the origial system) may have several solutions. Sometimes it is worth finding the solution with minimal Euclean norm.

Overdetermined linear systems

Proof:

$$F(x) = ||Ax - b||^2 = \langle A^*Ax, x\rangle - 2\langle x, A^*b\rangle + ||b||^2$$

The gradient vector of F at x is: $\operatorname{grad} F(x) = 2A^*Ax - 2A^*b$
At the location of minimum, the gradient vector vanishes.

$A^*A \in \mathbf{M}_{M\times M}$, therefore the size of the matrix of the Gaussian normal equation is much less than that of the original system, if $N \gg M$.

The Gaussian normal equation may be severely ill-conditioned.

The method of least squares remains applicable when $N < M$ (underdetermined systems). Now the Gaussian normal equation (and possibly also the origial system) may have several solutions. Sometimes it is worth finding the solution with minimal Euclidean norm.

Overdetermined linear systems

Proof:

$$F(x) = ||Ax - b||^2 = \langle A^*Ax, x\rangle - 2\langle x, A^*b\rangle + ||b||^2$$

The gradient vector of F at x is: $\operatorname{grad} F(x) = 2A^*Ax - 2A^*b$
At the location of minimum, the gradient vector vanishes.

$A^*A \in \mathbf{M}_{M\times M}$, therefore the size of the matrix of the Gaussian normal equation is much less than that of the original system, if $N \gg M$.

The Gaussian normal equation may be severely ill-conditioned.

The method of least squares remains applicable when $N < M$ (underdetermined systems). Now the Gaussian normal equation (and possibly also the origial system) may have several solutions. Sometimes it is worth finding the solution with minimal Eucliean norm.

Overdetermined linear systems

Proof:

$$F(x) = ||Ax - b||^2 = \langle A^*Ax, x\rangle - 2\langle x, A^*b\rangle + ||b||^2$$

The gradient vector of F at x is: $\operatorname{grad} F(x) = 2A^*Ax - 2A^*b$
At the location of minimum, the gradient vector vanishes.

$A^*A \in \mathbf{M}_{M\times M}$, therefore the size of the matrix of the Gaussian normal equation is much less than that of the original system, if $N \gg M$.

The Gaussian normal equation may be severely ill-conditioned.

The method of least squares remains applicable when $N < M$ (underdetermined systems). Now the Gaussian normal equation (and possibly also the origial system) may have several solutions. Sometimes it is worth finding the solution with minimal Euclean norm.

Approximation of functions by least squares

Let $f : [0,1] \to \mathbf{R}$ be a given function. Find a polynomial $p(x) = a_0 + a_1 x + ... + a_N x^N$ of degree at most N such that it approximates the function f 'as exactly as possible' on the interval $[0,1]$.

The exact fitting would be as follows:
$a_0 + a_1 x + ... + a_N x^N \equiv f(x)$
which does not hold in general. The 'best fitting' can be defined by e.g. minimizing the square integral of the residual:

$$F(\mathbf{a}) := \int_0^1 (a_0 + a_1 x + ... + a_N x^N - f(x))^2 \, dx \to \min!$$

Approximation of functions by least squares

Let $f : [0, 1] \to \mathbf{R}$ be a given function. Find a polynomial $p(x) = a_0 + a_1 x + ... + a_N x^N$ of degree at most N such that it approximates the function f 'as exactly as possible' on the interval $[0, 1]$.

The exact fitting would be as follows:
$a_0 + a_1 x + ... + a_N x^N \equiv f(x)$
which does not hold in general. The 'best fitting' can be defined by e.g. minimizing the square integral of the residual:

$$F(\mathbf{a}) := \int_0^1 (a_0 + a_1 x + ... + a_N x^N - f(x))^2 \, dx \to \min!$$

Approximation of functions by least squares

Let $f : [0,1] \to \mathbf{R}$ be a given function. Find a polynomial $p(x) = a_0 + a_1 x + ... + a_N x^N$ of degree at most N such that it approximates the function f 'as exactly as possible' on the interval $[0,1]$.

The exact fitting would be as follows:
$a_0 + a_1 x + ... + a_N x^N \equiv f(x)$
which does not hold in general. The 'best fitting' can be defined by e.g. minimizing the square integral of the residual:

$$F(\mathbf{a}) := \int_0^1 (a_0 + a_1 x + ... + a_N x^N - f(x))^2 \, dx \to \min !$$

Approximation of functions by least squares

Let $f : [0,1] \to \mathbf{R}$ be a given function. Find a polynomial $p(x) = a_0 + a_1 x + ... + a_N x^N$ of degree at most N such that it approximates the function f 'as exactly as possible' on the interval $[0,1]$.

The exact fitting would be as follows:
$a_0 + a_1 x + ... + a_N x^N \equiv f(x)$
which does not hold in general. The 'best fitting' can be defined by e.g. minimizing the square integral of the residual:

$$F(\mathbf{a}) := \int_0^1 (a_0 + a_1 x + ... + a_N x^N - f(x))^2 \, dx \to \min!$$

At the location of the minimum, the partial derivatives vanish:

$$\frac{\partial F}{\partial a_k} = \int_0^1 2 \cdot (a_0 + a_1 x + \ldots + a_N x^N - f(x)) \cdot x^k \, dx = 0$$

$$a_0 \int_0^1 x^k \, dx + a_1 \int_0^1 x^{k+1} \, dx + \ldots + a_N \int_0^1 x^{k+N} \, dx = \int_0^1 f(x) \cdot x^k \, dx$$

$(k = 0, 1, \ldots, N)$

The (k, j)th entry of the matrix of the system is: $A_{k,j} = \frac{1}{1+k+j}$ (**Hilbert matrix** of order $(N+1)$).

Approximation of functions by least squares

At the location of the minimum, the partial derivatives vanish:

$$\frac{\partial F}{\partial a_k} = \int_0^1 2 \cdot (a_0 + a_1 x + \ldots + a_N x^N - f(x)) \cdot x^k \, dx = 0$$

$$\boxed{a_0 \int_0^1 x^k \, dx + a_1 \int_0^1 x^{k+1} \, dx + \ldots + a_N \int_0^1 x^{k+N} \, dx = \int_0^1 f(x) \cdot x^k \, dx}$$

$(k = 0, 1, \ldots, N)$

The (k, j)th entry of the matrix of the system is: $A_{k,j} = \frac{1}{1+k+j}$ (**Hilbert matrix** of order $(N+1)$).

Approximation of functions by least squares

At the location of the minimum, the partial derivatives vanish:

$$\frac{\partial F}{\partial a_k} = \int_0^1 2 \cdot (a_0 + a_1 x + \ldots + a_N x^N - f(x)) \cdot x^k \, dx = 0$$

$$\boxed{a_0 \int_0^1 x^k \, dx + a_1 \int_0^1 x^{k+1} \, dx + \ldots + a_N \int_0^1 x^{k+N} \, dx = \int_0^1 f(x) \cdot x^k \, dx}$$

$(k = 0, 1, \ldots, N)$

The (k, j)th entry of the matrix of the system is: $A_{k,j} = \frac{1}{1+k+j}$ (**Hilbert matrix** of order $(N+1)$).

Approximation of functions by least squares

At the location of the minimum, the partial derivatives vanish:

$$\frac{\partial F}{\partial a_k} = \int_0^1 2 \cdot (a_0 + a_1 x + ... + a_N x^N - f(x)) \cdot x^k \, dx = 0$$

$$\boxed{a_0 \int_0^1 x^k \, dx + a_1 \int_0^1 x^{k+1} \, dx + ... + a_N \int_0^1 x^{k+N} \, dx = \int_0^1 f(x) \cdot x^k \, dx}$$

$(k = 0, 1, ..., N)$

The (k,j)th entry of the matrix of the system is: $A_{k,j} = \frac{1}{1+k+j}$ (**Hilbert matrix** of order $(N+1)$).

Least squares approximation of functions, example

Example: Approximate the function

$$f(x) := \cos\frac{\pi x}{2} \qquad (x \in [-1,\, 1])$$

by an at most second-order polynomial:

$$p(x) := a_0 + a_1 x + a_2 x^2$$

Least squares approximation of functions, example

1a) Solution by quadratic regression:

Consider the base points: $M := 3$, $x_1 := -1$, $x_2 := 0$, $x_3 := 1$.
The associated values: $f_1 = 0$, $f_2 = 1$, $f_3 = 0$.
The system of equations for the regression coefficients:

$$\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\rightarrow \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

The regression polynomial: $p(x) = a_0 + a_1 x + a_2 x^2 = 1 - x^2$.

Least squares approximation of functions, example

1a) Solution by quadratic regression:

Consider the base points: $M := 3,\ x_1 := -1,\ x_2 := 0,\ x_3 := 1.$

The associated values: $f_1 = 0,\ f_2 = 1,\ f_3 = 0.$

The system of equations for the regression coefficients:

$$\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\to \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

The regression polynomial: $p(x) = a_0 + a_1 x + a_2 x^2 = 1 - x^2.$

Least squares approximation of functions, example

1a) Solution by quadratic regression:

Consider the base points: $M := 3$, $x_1 := -1$, $x_2 := 0$, $x_3 := 1$.
The associated values: $f_1 = 0$, $f_2 = 1$, $f_3 = 0$.
The system of equations for the regression coefficients:

$$\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\rightarrow \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

The regression polynomial: $p(x) = a_0 + a_1 x + a_2 x^2 = 1 - x^2$.

Least squares approximation of functions, example

1a) Solution by quadratic regression:

Consider the base points: $M := 3$, $x_1 := -1$, $x_2 := 0$, $x_3 := 1$.
The associated values: $f_1 = 0$, $f_2 = 1$, $f_3 = 0$.
The system of equations for the regression coefficients:

$$\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\rightarrow \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

The regression polynomial: $p(x) = a_0 + a_1 x + a_2 x^2 = 1 - x^2$.

Least squares approximation of functions, example

1a) Solution by quadratic regression:

Consider the base points: $M := 3$, $x_1 := -1$, $x_2 := 0$, $x_3 := 1$.
The associated values: $f_1 = 0$, $f_2 = 1$, $f_3 = 0$.
The system of equations for the regression coefficients:

$$\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\rightarrow \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

The regression polynomial: $p(x) = a_0 + a_1 x + a_2 x^2 = 1 - x^2$.

Least squares approximation of functions, example

1a) Solution by quadratic regression:

Consider the base points: $M := 3$, $x_1 := -1$, $x_2 := 0$, $x_3 := 1$.
The associated values: $f_1 = 0$, $f_2 = 1$, $f_3 = 0$.
The system of equations for the regression coefficients:

$$\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$\to \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

The regression polynomial: $p(x) = a_0 + a_1 x + a_2 x^2 = 1 - x^2$.

Least squares approximation of functions, example

1b) Solution by quadratic regression. Now define $M := 5$, and:

$x_1 := -1$, $x_2 := -0.5$, $x_3 := 0$, $x_4 := 0.5$, $x_5 := 1$.
The associated values: $f_1 = 0$, $f_2 = \frac{\sqrt{2}}{2}$, $f_3 = 1$, $f_4 = \frac{\sqrt{2}}{2}$, $f_5 = 0$.
The system of equations for the regression coefficients:

$$\begin{pmatrix} 5 & 0 & 2.5 \\ 0 & 2.5 & 0 \\ 2.5 & 0 & 2.125 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1+\sqrt{2} \\ 0 \\ \frac{\sqrt{2}}{4} \end{pmatrix} = \begin{pmatrix} 2.41421 \\ 0 \\ 0.35355 \end{pmatrix}$$

$$\rightarrow \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.97059 \\ 0 \\ -0.97549 \end{pmatrix}$$

The regression polynomial: $p(x) = 0.97059 - 0.97549x^2$.

Least squares approximation of functions, example

1b) Solution by quadratic regression. Now define $M := 5$, and:

$x_1 := -1$, $x_2 := -0.5$, $x_3 := 0$, $x_4 := 0.5$, $x_5 := 1$.

The associated values: $f_1 = 0$, $f_2 = \frac{\sqrt{2}}{2}$, $f_3 = 1$, $f_4 = \frac{\sqrt{2}}{2}$, $f_5 = 0$.

The system of equations for the regression coefficients:

$$\begin{pmatrix} 5 & 0 & 2.5 \\ 0 & 2.5 & 0 \\ 2.5 & 0 & 2.125 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1+\sqrt{2} \\ 0 \\ \frac{\sqrt{2}}{4} \end{pmatrix} = \begin{pmatrix} 2.41421 \\ 0 \\ 0.35355 \end{pmatrix}$$

$$\to \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.97059 \\ 0 \\ -0.97549 \end{pmatrix}$$

The regression polynomial: $p(x) = 0.97059 - 0.97549x^2$.

Least squares approximation of functions, example

1b) Solution by quadratic regression. Now define $M := 5$, and:

$x_1 := -1$, $x_2 := -0.5$, $x_3 := 0$, $x_4 := 0.5$, $x_5 := 1$.
The associated values: $f_1 = 0$, $f_2 = \frac{\sqrt{2}}{2}$, $f_3 = 1$, $f_4 = \frac{\sqrt{2}}{2}$, $f_5 = 0$.

The system of equations for the regression coefficients:

$$\begin{pmatrix} 5 & 0 & 2.5 \\ 0 & 2.5 & 0 \\ 2.5 & 0 & 2.125 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1+\sqrt{2} \\ 0 \\ \frac{\sqrt{2}}{4} \end{pmatrix} = \begin{pmatrix} 2.41421 \\ 0 \\ 0.35355 \end{pmatrix}$$

$$\rightarrow \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.97059 \\ 0 \\ -0.97549 \end{pmatrix}$$

The regression polynomial: $p(x) = 0.97059 - 0.97549x^2$.

Least squares approximation of functions, example

1b) Solution by quadratic regression. Now define $M := 5$, and:

$x_1 := -1$, $x_2 := -0.5$, $x_3 := 0$, $x_4 := 0.5$, $x_5 := 1$.
The associated values: $f_1 = 0$, $f_2 = \frac{\sqrt{2}}{2}$, $f_3 = 1$, $f_4 = \frac{\sqrt{2}}{2}$, $f_5 = 0$.
The system of equations for the regression coefficients:

$$\begin{pmatrix} 5 & 0 & 2.5 \\ 0 & 2.5 & 0 \\ 2.5 & 0 & 2.125 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1+\sqrt{2} \\ 0 \\ \frac{\sqrt{2}}{4} \end{pmatrix} = \begin{pmatrix} 2.41421 \\ 0 \\ 0.35355 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.97059 \\ 0 \\ -0.97549 \end{pmatrix}$$

The regression polynomial: $p(x) = 0.97059 - 0.97549x^2$.

Least squares approximation of functions, example

1b) Solution by quadratic regression. Now define $M := 5$, and:

$x_1 := -1,\ x_2 := -0.5,\ x_3 := 0,\ x_4 := 0.5,\ x_5 := 1.$
The associated values: $f_1 = 0,\ f_2 = \frac{\sqrt{2}}{2},\ f_3 = 1,\ f_4 = \frac{\sqrt{2}}{2},$
$f_5 = 0.$
The system of equations for the regression coefficients:

$$\begin{pmatrix} 5 & 0 & 2.5 \\ 0 & 2.5 & 0 \\ 2.5 & 0 & 2.125 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1+\sqrt{2} \\ 0 \\ \frac{\sqrt{2}}{4} \end{pmatrix} = \begin{pmatrix} 2.41421 \\ 0 \\ 0.35355 \end{pmatrix}$$

$$\to \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.97059 \\ 0 \\ -0.97549 \end{pmatrix}$$

The regression polynomial: $p(x) = 0.97059 - 0.97549x^2.$

1b) Solution by quadratic regression. Now define $M := 5$, and:

$x_1 := -1,\ x_2 := -0.5,\ x_3 := 0,\ x_4 := 0.5,\ x_5 := 1.$

The associated values: $f_1 = 0,\ f_2 = \frac{\sqrt{2}}{2},\ f_3 = 1,\ f_4 = \frac{\sqrt{2}}{2},$ $f_5 = 0.$

The system of equations for the regression coefficients:

$$\begin{pmatrix} 5 & 0 & 2.5 \\ 0 & 2.5 & 0 \\ 2.5 & 0 & 2.125 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1+\sqrt{2} \\ 0 \\ \frac{\sqrt{2}}{4} \end{pmatrix} = \begin{pmatrix} 2.41421 \\ 0 \\ 0.35355 \end{pmatrix}$$

$$\rightarrow \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.97059 \\ 0 \\ -0.97549 \end{pmatrix}$$

The regression polynomial: $p(x) = 0.97059 - 0.97549x^2$.

Least squares approximation of functions, example

2) Solution by minimizing the difference measured in L_2-norm:

$$\int_{-1}^{1} (a_0 + a_1 x + a_2 x^2 - f(x))^2 \, dx \to \min!$$

$$\begin{pmatrix} \int_{-1}^{1} 1 \, dx & \int_{-1}^{1} x \, dx & \int_{-1}^{1} x^2 \, dx \\ \int_{-1}^{1} x \, dx & \int_{-1}^{1} x^2 \, dx & \int_{-1}^{1} x^3 \, dx \\ \int_{-1}^{1} x^2 \, dx & \int_{-1}^{1} x^3 \, dx & \int_{-1}^{1} x^4 \, dx \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \int_{-1}^{1} f(x) \, dx \\ \int_{-1}^{1} f(x) \cdot x \, dx \\ \int_{-1}^{1} f(x) \cdot x^2 \, dx \end{pmatrix}$$

Least squares approximation of functions, example

2) Solution by minimizing the difference measured in L_2-norm:

$$\int_{-1}^{1} (a_0 + a_1 x + a_2 x^2 - f(x))^2 \, dx \to \min!$$

$$\begin{pmatrix} \int_{-1}^{1} 1 \, dx & \int_{-1}^{1} x \, dx & \int_{-1}^{1} x^2 \, dx \\ \int_{-1}^{1} x \, dx & \int_{-1}^{1} x^2 \, dx & \int_{-1}^{1} x^3 \, dx \\ \int_{-1}^{1} x^2 \, dx & \int_{-1}^{1} x^3 \, dx & \int_{-1}^{1} x^4 \, dx \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \int_{-1}^{1} f(x) \, dx \\ \int_{-1}^{1} f(x) \cdot x \, dx \\ \int_{-1}^{1} f(x) \cdot x^2 \, dx \end{pmatrix}$$

Least squares approximation of functions, example

2) Solution by minimizing the difference measured in L_2-norm:

$$\int_{-1}^{1} (a_0 + a_1 x + a_2 x^2 - f(x))^2 \, dx \to \min!$$

$$\begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1.27324 \\ 0 \\ 0.24119 \end{pmatrix}$$

$$\to \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.98016 \\ 0 \\ -1.03063 \end{pmatrix}$$

The best approximation:
$p(x) = a_0 + a_1 x + a_2 x^2 = 0.98016 - 1.03063x^2$.

Least squares approximation of functions, example

2) Solution by minimizing the difference measured in L_2-norm:

$$\int_{-1}^{1} (a_0 + a_1 x + a_2 x^2 - f(x))^2 \, dx \to \min!$$

$$\begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1.27324 \\ 0 \\ 0.24119 \end{pmatrix}$$

$$\to \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.98016 \\ 0 \\ -1.03063 \end{pmatrix}$$

The best approximation:
$p(x) = a_0 + a_1 x + a_2 x^2 = 0.98016 - 1.03063 x^2$.

2) Solution by minimizing the difference measured in L_2-norm:

$$\int_{-1}^{1} (a_0 + a_1 x + a_2 x^2 - f(x))^2 \, dx \to \min!$$

$$\begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 1.27324 \\ 0 \\ 0.24119 \end{pmatrix}$$

$$\to \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0.98016 \\ 0 \\ -1.03063 \end{pmatrix}$$

The best approximation:
$p(x) = a_0 + a_1 x + a_2 x^2 = 0.98016 - 1.03063 x^2$.