Least Squares Approximation

Marieke and Eren

21.05.2024

Problem

- Find a best fit polynomial of degree n for $m+1 \in \mathbb{N}$ points $(x_i, y_i)_{0 \leq i \leq m}$ in $\mathbb{R}^2$
- Solution: Find the vector $(a_0, ...a_n) \in \mathbb{R}^{n+1}$ that minimizes $(y_0 - (a_0 + a_1 x_0 + ... + a_n x_0^n))^2 + ... + (y_m - (a_0 + a_1 x_m + ... + a_n x_m^n))^2$

Transition to linear algebra

Let $Y = \begin{pmatrix} y_0 \\ . \\ . \\ . \\ y_m \end{pmatrix} \in \mathbb{R}^{m+1}$, $X = \begin{pmatrix} a_n \\ . \\ . \\ . \\ a_0 \end{pmatrix} \in \mathbb{R}^{n+1}$ and

$A = \begin{pmatrix} x_0^n & ... & x_0 & 1 \\ . & ... & . & . \\ . & ... & . & . \\ . & ... & . & . \\ x_m^n & ... & x_m & 1 \end{pmatrix} \in \mathbb{R}^{m+1 \times n+1}$. Now find $\min_{X \in \mathbb{R}^{n+1}} \|Y - AX\|^2$.

Theorem 1

Let $Y \in \mathbb{R}^k$ for some $k \in \mathbb{N}$ and U a subspace of $\mathbb{R}^k$. Let $Y = Y_U + Y_{U^\perp}$ be the orthogonal decomposition of Y with regards to U. Then
$\|Y - Y_U\| \leq \|Y - X\|$
for all $X \in U$. The inequality is an equality if and only if $X = Y_U$.

Proof

Let $X \in U$. Then
$\|Y - Y_U\|^2 \leq \|Y - Y_U\|^2 + \|Y_U - X\|^2 = \|Y - Y_U + Y_U - X\|^2 = \|Y - X\|^2$
by the pythagorean theorem.

Theorem 2

Let $k \leq n$ and let $A \in \mathrm{Mat}(n \times k; \mathbb{R})$ be a matrix, which we think of as a family of k vectors in $\mathbb{R}^n$. Let P be the matrix, in the canonical basis of $\mathbb{R}^n$, of the orthogonal projection to $\mathrm{Im} A$. If $\mathrm{rank} A = k$, then the matrix $A^t A \in Mat(n \times n; \mathbb{R})$ is invertible and we have $P = A(A^t A)^1 A^t$.

Proof - A^tA is invertible

- A full rank $\implies$ columns of A form basis of $\mathrm{Im}A$
  - Note: $\dim(\mathrm{Im}A) = k$
- A^tA invertible $\iff$ left multiplication by A^tA bijective
- Let $Y \in \mathrm{Ker}A^tA$. Then $0 = Y^tA^tAY = (AY)^tAY = \|AY\|^2$ $\implies AY = 0$. By rank-nullity-theorem we know $\dim(\mathrm{Ker}A) = 0 \implies Y = 0 \implies \mathrm{Ker}A^tA = 0$ $\implies \dim(\mathrm{Ker}A^tA) = 0$. By rank-nullity theorem we obtain $\dim(\mathrm{Im}A^tA) = n$.

Proof - Calculating P

Let $Y \in \mathbb{R}^k$ and $Y = Y_{\mathrm{Im}A} + Y_{(\mathrm{Im}A)^\perp}$ be its orthogonal decomposition.

- $Y_{\mathrm{Im}A} \in \mathrm{Im}A \implies \exists X \in \mathbb{R}^n : AX = Y_{\mathrm{Im}A} \implies Y_{(\mathrm{Im}A)^\perp} = Y - Y_{\mathrm{Im}A} = Y - AX \in (\mathrm{Im}A)^\perp$
- rank-nullity theorem $\implies \dim(\mathrm{Im}A) + \dim(\mathrm{Im}A)^\perp = n = \dim(\mathrm{Ker}A^t) + \dim(\mathrm{Im}A^t) = \dim(\mathrm{Ker}A^t) + \dim(\mathrm{Im}A) \implies \dim(\mathrm{Im}A)^\perp = \dim(\mathrm{Ker}A^t) \implies (\mathrm{Im}A)^\perp = \mathrm{Ker}A^t$
- $\implies 0 = A^t(Y - AX) = A^tY - A^tAX \implies A^tY = A^tAX \implies (A^tA)^{-1}A^tY = X \implies A(A^tA)^{-1}A^tY = AX = Y_{\mathrm{Im}A}$
- $\implies P = A(A^tA)^{-1}A^t$.

Corollary 3

Let $k \leq n$ and let $A \in \mathrm{Mat}(n \times k; \mathbb{R})$ be a matrix. If $\mathrm{rank} A = k$, then for all $Y \in \mathbb{R}^n$, the least squares minimisation problem $\min_{X \in \mathbb{R}^k} \|Y - AX\|^2$ admits the vector $X = (A^t A)^{-1} A^t Y$ as its unique solution.

Proof

$\|Y - AX\|$ is minimal for $AX = Y_{\mathrm{Im}A} = PY$. Then $X = (A^tA)^{-1}A^tY$, from Theorem 2.

Corollary 4

Let $A \in \mathrm{Mat}(n \times n; \mathbb{R})$ be a matrix. If $\mathrm{rank}(A) = n$, then A is invertible and for all $Y \in \mathbb{R}^n$, the least squares minimisation problem $\min_{X \in \mathbb{R}^n} \|Y - AX\|^2$ admits the vector $X = A^{-1}Y$ as its unique solution.

Proof

A has full rank $\implies$ A is invertible. Corollary 3
$\implies X = (A^t A)^{-1} A^t Y = A^{-1} (A^t)^{-1} A^t Y = A^{-1} Y$

Problem

- Find a polynomial $L(T) \in \mathbb{R}[T]$ of $\deg L \leq m$ that passes through each data point in $(x_i, y_i)_{0 \leq i \leq m}$, so $L(x_i) = y_i$ for each x_i.
- Solution: Solve the system of linear equations $\sum_{i=0}^{m} a_i x_k^i = y_k, 0 \leq k \leq m$ where $(a_i)_{0 \leq i \leq m}$ are the coefficients of $L(T)$ in standard basis.

Transition to linear algebra

Let $Y = \begin{pmatrix} y_0 \\ . \\ . \\ . \\ y_m \end{pmatrix} \in \mathbb{R}^{m+1}$, $X = \begin{pmatrix} a_m \\ . \\ . \\ . \\ a_0 \end{pmatrix} \in \mathbb{R}^{m+1}$ and

$$A = \begin{pmatrix} x_0^m & ... & x_0 & 1 \\ . & ... & . & . \\ . & ... & . & . \\ . & ... & . & . \\ x_m^m & ... & x_m & 1 \end{pmatrix} \in \mathbb{R}^{m+1 \times m+1}.$$ Then $AX = Y$.

Comparison with Corollary 4 $\implies$ If we use least squares method to approximate by a function of $\deg = m$ the polynomial we obtain coincides with the Lagrange polynomial.

Theorem 5

For a set of points $(x_i, y_i)_{0 \leq i \leq n} \subseteq \mathbb{R}^2$ with $x_i \neq x_j$ for $i \neq j$ let $L(x) \in \mathbb{R}[x]$, $\deg L = n$, be defined as $L(x) = \sum_{i=0}^{n} y_i l_i(x)$ for $l_i(x) = \prod_{0 \leq j \leq n, j \neq i} \frac{x - x_j}{x_i - x_j}$. Then $L(x)$ is unique and it fulfills $L(x_i) = y_i$ for each i.

Proof - Basis

- Note: $l_i(x_k) = 0$ for $k \neq i$ and $l_i(x_i) = 1$.
- Show: $\{l_i(x)\}_{0 \leq i \leq n}$ form a generating set of V the vector space of polynomials of degree less than or equal to n over $\mathbb{R}$.
- $\dim V = n + 1$.
- Let $f(x) \in V$. Def $q(x) := f(x) - \sum_{i=0}^{n} f(x_i) l_i(x)$
- $\deg l_i = n$ for each i and $\deg f \leq n \implies \deg q \leq n$
- $q(x_i) = 0$ for each $i \implies q$ has $n + 1$ roots
  $\implies q = 0 \implies f(x) = \sum_{i=0}^{n} f(x_i) l_i(x)$
- dimensional reasons $\implies \{l_i(x)\}_{0 \leq i \leq n}$ form a basis of V.

Proof - Uniqueness

- In particular: $\sum_{i=0}^{n} y_i l_i(x) = \sum_{i=0}^{n} L(x_i) l_i(x)$ is a unique representation of $L(x)$.
- $L(x)$ is unique: Let $P(x) \in V$ be a polynomial with $P(x_i) = y_i$ for each $i \implies P(x) = \sum_{i=0}^{n} P(x_i) l_i(x) = \sum_{i=0}^{n} y_i l_i(x) = \sum_{i=0}^{n} L(x_i) l_i(x) = L(x)$