Approximation of functions and data

by a simpler and convenient representation

Two core approaches:

1. Interpolation: for reliable data, want to reproduce the values (exact)

2. Least Squares fit: for unreliable data (that may contain errors) want to capture the trend of data (not the values).

Data points (x_i, y_i), $i=1,\dots,m$ should be thought of as values $y_i = f(x_i)$ of some function $y=f(x)$ (known or unknown)

Want to approximate / represent $y=f(x)$ by a simple/convenient model $y = \Phi(x)$.

Approximation of $f(x)$ based on:

a. values at a single point x_0: Taylor polynomial $P_k(x)$ of degree k if $f \in C^{k+1}$: matches $f^{(i)}$ only at x_0

Padé (rational fn) $\dfrac{P_m(x)}{Q_n(x)}$: matches $f^{(i)}$ only at x_0

Chebyshev rational approx. $\dfrac{\sum p_i T_i(x)}{\sum q_i T_i(x)}$, $T_i(x)$ = Chebyshev poly of degree i

b. values at several points $\{x_i\}$ (nodes):

Interpolating poly. $P_N(x)$ at $N+1$ distinct nodes: matches f at the nodes $\{x_i\}$

Hermite interpolating poly: matches f, f' at the nodes

Cubic spline : matches only f at the nodes

piecewise interpolant : " " f " " "

Least Squares fit to a model Φ

c. values on an interval:

Trig polynomials (Fourier expansion): Inverse Discrete Fourier Transform

Wavelets

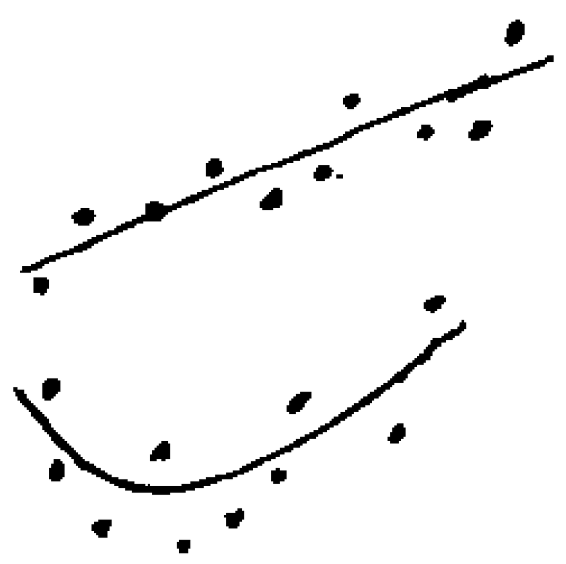
Least Squares fit to a model Φ

d. values on $(-\infty, \infty)$: Integral Fourier Transform

Least Squares fit of a model to data (x_i, y_i), $i=1:m$

(Regression in Statistics)

When data may contain errors (in measurement, etc), interpolation is not appropriate.



Want a (simple) representation that captures the trend:

Find a model $y=\Phi(x)$ that "best fits" the data. $y_i = f(x_i)$.

"best" in what sense? need a measure of being "close", of least error.

Gauss proposed the "Least Squares" sense:

Least sum of squared deviations:

Find coefficients $\vec{c}$ in a model $y=\Phi(x, \vec{c})$ to minimize the

Least Squares error $\quad E(\vec{c}) = \sum_{i=1}^{m} \left(y_i - \Phi(x_i, \vec{c}) \right)^2$

e.g. fit a line $y = ax+b$: Find a,b to $\min\limits_{a,b} \sum_{i=1}^{m} \left(y_i - (ax_i + b) \right)^2$

Can be done by Calculus: $\nabla E(a,b) = \vec{0}$ i.e. $\frac{\partial E}{\partial a} = 0$, $\frac{\partial E}{\partial b} = 0$:

$$\begin{cases} \frac{\partial E}{\partial a} = \sum_{1}^{m} 2\left(y_i - [ax_i + b]\right)\cdot(-x_i) = 0 \\ \frac{\partial E}{\partial b} = \sum_{1}^{m} 2\left(\qquad \right)\cdot(-1) = 0 \end{cases} \Rightarrow \begin{cases} a\sum x_i^2 + b\sum x_i = \sum x_i y_i \\ a\sum x_i + b\cdot m = \sum y_i \end{cases}$$

normal equations

$$\Rightarrow \quad a = \frac{m\sum x_i y_i - \left(\sum x_i\right)\left(\sum y_i\right)}{m\sum x_i^2 - \left(\sum x_i\right)^2}, \quad b = \frac{\left(\sum x_i^2\right)\left(\sum y_i\right) - \left(\sum x_i y_i\right)\left(\sum x_i\right)}{m\sum x_i^2 - \left(\sum x_i\right)^2}$$

the regression line

Least Squares error = R-square error = $\min E(a,b) \geq 0$

R^2 error

Common choices for model $y = \Phi(x, \vec{c})$: polynomial $P_n(x)$, $n < m$

$a e^{bx}$, $a \cos \omega x + b \sin \omega x$, ...

Linear Least Squares: fit to a "linear model": $\Phi(x, \vec{c}) = \sum_{k=1}^{n} c_k \cdot \varphi_k(x)$,

i.e. to a linear combination of basis functions $\{\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)\}$.
$n \ll m$

e.g. choosing the powers x^k as basis functions $\varphi_k(x) = x^k$, $k = 0, 1, \dots, n-1$
amounts to fitting an n-th degree polynomial (and leads to ill-conditioned system!)

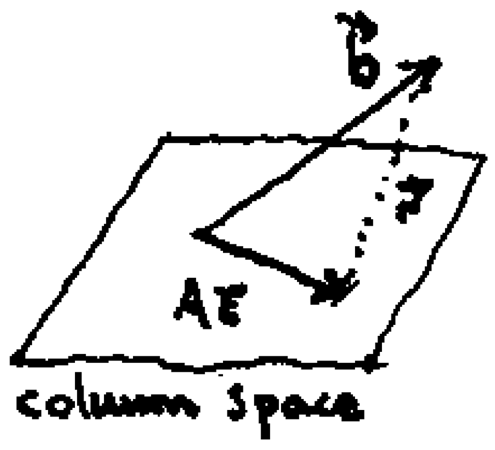
LS error: $$E(\vec{c}) = \sum_{i=1}^{m} \Big(y_i - \sum_{k=1}^{n} c_k \varphi_k(x_i) \Big)^2$$

LS problem: Find coefficients $\vec{c} = (c_1, \dots, c_n)$ to minimize $E(\vec{c})$.

Set $A = [a_{ik}] = [\varphi_k(x_i)]$ $m \times n$ matrix, $m \geq n$ (usually $m \gg n$).
$\vec{b} = \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$ m-vector

then $E(\vec{c}) = \|\vec{b} - A\vec{c}\|_2^2 = \|\text{residual}\|_2^2$ for the system $A\vec{c} \text{ "="} \vec{b}$

which is overdetermined (m : equs, m, than unknowns, n) when $m > n$,
so it has no solution in general!



All we want is a vector $A\vec{c}$ closest to $\vec{b}$ (in Euclidean 2-norm)
i.e. smallest possible residual $\vec{r} = \vec{b} - A\vec{c}$
which is the orthogonal projection of $\vec{b}$ onto the column space of A.
$\vec{r} \perp \text{col}(A) \Leftrightarrow \vec{r} \in N(A^T) \Leftrightarrow A^T\vec{r} = \vec{0} \Leftrightarrow A^T(\vec{b} - A\vec{c}) = \vec{0} \Leftrightarrow$

$$A^T A \vec{c} = A^T \vec{b} \quad : \text{ the normal equations for } \vec{c}$$
(same as $\nabla E(\vec{c}) = \vec{0}$!)

This is $n \times n$ linear system for $\vec{c}$, ~~and~~ $A^T A$ is symmetric $n \times n$,
but often ill-conditioned: $\kappa(A^T A) = (\kappa(A))^2$, <u>not</u> for Gauss elim. methods!!!
and if $\text{rank}(A) < n$ then $A^T A$ is singular, system has ∞-ly many solutions.

Backslash operator of Matlab: $x = A \backslash b$ returns a LS-sense solution of $Ax = b$, any A $m \times n$,
i.e. finds x to $\min_x \|b - Ax\|_2^2$, via QR factorization of A:

$A = Q\begin{bmatrix} R \\ 0 \end{bmatrix}$, Q $m \times m$ orthogonal ($Q^T Q = I$), R upper triangular $n \times n$

Then solve upper triangular $Rx = \left[\begin{matrix} Q^T b \\ 0 \end{matrix}\right]\}$ n rows by back substitutions!
The magic is orthogonality!!!

The space $L^2(a,b) = \left\{ f : (a,b) \to \underset{\mathbb{C}}{\overset{}{\mathbb{R}}}\ \text{or}\ \mathbb{C} \text{ with } \int_a^b |f(x)|^2 dx < \infty \right\}$

= space of all square-integrable functions defined on (a,b).

It is a vector space, a normed space with norm $\|f\|_{L^2} = \sqrt{\int_a^b |f(x)|^2 dx}$,

and inner-product space with inner product $\langle f, g \rangle = \int_a^b f(x) g(x) dx$.

$\langle f, f \rangle = \|f\|^2$, so $\|f\| = \sqrt{\langle f, f \rangle}$.

These generalize the 2-norm and dot product of vectors in $\mathbb{R}^n$. to infinite dimensions

Orthogonality: $f \perp g$ iff $\langle f, g \rangle = 0$

Orthogonal set: $\{\varphi_k\}_{k=1,\dots,n}$ in L^2 if $\langle \varphi_k, \varphi_\ell \rangle = 0$ for $k \neq \ell$

Orthonormal set: $\{\varphi_k\}_{k=1:n}$ in L^2 if $\langle \varphi_k, \varphi_\ell \rangle = \delta_{k\ell} = \begin{cases} 0, & k \neq \ell \\ 1, & k = \ell \end{cases}$

Complete orthonormal set $\{\varphi_k\}_{k=1,2,\dots}$ in L^2 if $\langle f, \varphi_k \rangle = 0\ \forall k \Rightarrow f = 0$
$\equiv$ orthonormal basis for L^2.

Weighted $L^2_w(a,b) = \left\{ f : (a,b) \to \mathbb{R},\ \|f\|^2_{L^2_w} = \int_a^b w(x) |f(x)|^2 dx \right\}$, $\langle f, g \rangle_{L^2_w} = \int_a^b w(x) f(x) g(x) dx$
$w(x) > 0$

Most important orthogonal bases:

1. Orthogonal polynomials; for $L^2_w(a,b)$
   - Legendre for $(-1,1)$, $w(x) \equiv 1$ $\quad \{1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x, \dots\}$
   - Chebyshev for $(-1,1)$ with $w(x) = \frac{1}{\sqrt{1-x^2}}$ $\quad \{1, x, 2x^2-1, 4x^3-3x, \dots\}$
   - Laguerre for $(0,\infty)$ with $w(x) = e^{-x}$ $\quad \{1, 1-x, \frac{1}{2}(x^2-4x+2), \dots\}$
   - Hermite for $(-\infty,\infty)$ with $w(x) = e^{-x^2}$ $\quad \{1, 2x, 4x^2-2, 8x^3-12x, \dots\}$
2. In $L^2(0,L)$: $\left\{ \cos \frac{k\pi x}{L} : k = 0, 1, \dots \right\}$, and $\left\{ \sin \frac{k\pi x}{L} : k = 1, 2, \dots \right\}$
3. In $L^2(-L,L)$: $\left\{ 1, \cos \frac{k\pi x}{L}, \sin \frac{k\pi x}{L} : k = 1, 2, \dots \right\}$
4. Eigenfunctions of self-adjoint Sturm-Liouville problems
5. Orthogonal wavelets of various types, e.g. Daubechies wavelets

Note: orthogonal set $\Rightarrow$ linearly independent set
so an $\perp$ set is a basis for its span (= all linear combinations of the set)

Gram-Schmidt orthogonalization

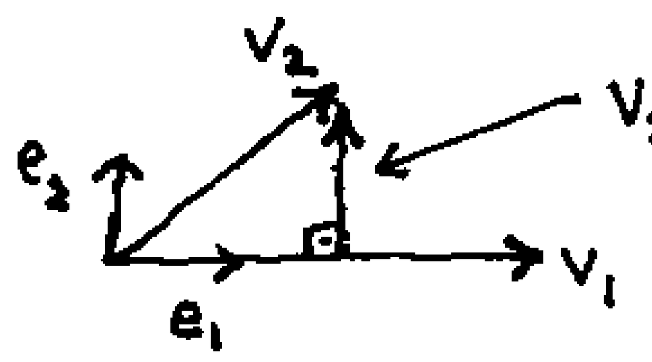
In any inner product space, e.g. $\mathbb{R}^n$, $\mathbb{C}^n$, L^2, L^2_w, ...

from a linearly independent set $\{v_1, v_2, \dots, v_N\}$ we can construct

an orthogonal or orthonormal set $\{e_1, e_2, \dots, e_N\}$ which spans the same subspace

$$e_1 = \frac{v_1}{\|v_1\|}$$

$$e_2 = \frac{v_2 - \langle v_2, e_1\rangle e_1}{\| \qquad\qquad \|}$$



$v_2 - \langle v_2, e_1\rangle e_1 \perp e_1$

$$e_3 = \frac{v_3 - \langle v_3, e_1\rangle e_1 - \langle v_3, e_2\rangle e_2}{\| \qquad\qquad\qquad \|}$$

$$\langle v_2, e_1\rangle - \langle v_2, e_1\rangle \underbrace{\langle e_1, e_1\rangle}_{\|e_1\|^2=1} = 0 \quad \therefore\ e_2 \perp e_1$$

...

Since e_i's are linear combinations of v_i's, they have same span.

Applying Gram-Schmidt to the linearly independent set $\{1, x, x^2, \dots\}$

in $L^2_w(a,b)$ we get the standard sets of orthogonal polynomials:

Legendre in $L^2(-1,1)$, Chebyshev in $L^2_w(-1,1)$, $w = \frac{1}{\sqrt{1-x^2}}$, Laguerre in $L^2_w(0,\infty)$ $w = e^{-x}$, Hermite in $L^2_w(-\infty,\infty)$ $w = e^{-x^2}$

(or any other (a,b) and weight $w(x) > 0$ we may want).

Each orthogonal / orthonormal set $\{\varphi_k\}$ in $L^2_w(a,b)$ constitutes a basis for $L^2_w(a,b)$.

Therefore any $f \in L^2_w$ has unique expansion $f \overset{L^2_w}{=} \sum_{k=1}^{\infty} c_k \varphi_k$ with $c_k = \frac{\langle f, \varphi_k\rangle}{\langle \varphi_k, \varphi_k\rangle}$.

Fundamental Theorem on (generalized) Fourier expansions:

For an orthonormal set $\{\varphi_k\}$ in a Hilbert space $\mathcal{H}$, the following are equivalent:

1. $\{\varphi_k\}$ is complete (constitutes an ON basis for $\mathcal{H}$).
2. Each $f \in \mathcal{H}$ has unique Fourier expansion w.r.t. $\{\varphi_k\}$: $f \overset{\mathcal{H}}{=} \sum c_k \varphi_k$, $c_k = \langle f, \varphi_k\rangle$
3. For each $f \in \mathcal{H}$, $\sum |c_k|^2 = \|f\|^2$ (Parceval equality, Pythagorean Thm).

If $\{\varphi_k\}$ is finite set, then $\mathcal{H}$ is finite dimensional (isomorphic to $\mathbb{R}^n$ or $\mathbb{C}^n$).

Else $\mathcal{H}$ is an infinite dimensional separable Hilbert space, like $L^2_w(a,b)$.