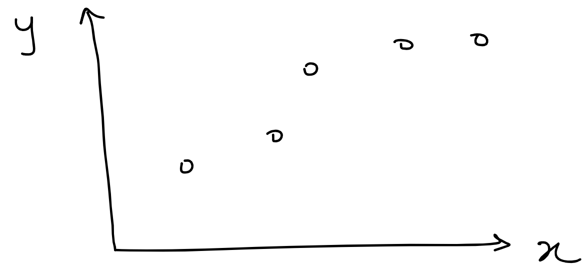
Function Approximation [Curve Fitting]

Regression [linear or ordinary least squares regression]

Given (x_i, y_i) $i = 1, 2, \ldots, N$



Objective - To fit a curve [polynomial]

Model - $\hat{y}_i = \sum_{j=0}^{m} a_j \phi_j(x_i) = \Phi a$

ϕ_j are basis functions

Design matrix $\underset{N \times m+1}{\Phi} = [\Phi_0 \; \Phi_1 \; \Phi_2 \; \ldots \; \Phi_m]$

$$\Phi_j = \begin{bmatrix} \phi_j(x_1) \\ \phi_j(x_2) \\ \vdots \\ \phi_j(x_N) \end{bmatrix}$$

Example - Quadratic polynomial

$$\Phi = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix}$$

$$\Phi_0 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \Phi_1 = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\Phi_2 = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{bmatrix}$$

Unknown - Regression coefficients

$$a = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}$$

Method - (optimization)

Principle of Least Squares

minimize $\sum e_i^2 = \min \sum (y_i - \hat{y}_i)^2$

Solution Normal equations

$$\Phi^T \Phi a = \Phi^T Y \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

System of 'm+1' linear equations

Problem - $\Phi^T \Phi$ has a large condition number

To avoid this problem

Select orthogonal basis functions so that the matrix $\Phi^T \Phi$ becomes diagonal

Orthogonal functions -

A set of functions $f_i(t)$ is said to be orthogonal w.r.t weigh vector $w(t)$ and over the interval $a \le x \le b$ if

$$\langle f_i f_j \rangle = \int_a^b w(t) f_i(t) f_j(t)\, dt = \begin{cases} 0 & i \ne j \\ \lambda_i > 0 & i = j \end{cases}$$

If $\int_a^b w(t) f_i^2(t)\, dt = 1$

the system is called Orthonormal

Orthogonal Polynomials

Gram-Schmidt Process – Given a set of independent functions $f_0, f_1, \ldots f_m$ a set of orthogonal functions $\theta_0, \theta_1, \theta_2, \ldots \theta_m$ [in the same subspace] are obtained as

$$\theta_0 = \frac{f_0}{\|f_0\|} \qquad \|f_0\| = \sqrt{\int_a^b f_0^2(t)\, dt}$$

$$\langle \theta_0 f_1 \rangle = \int_a^b \theta_0 f_1 \, dt$$

$$f_1' = f_1 - \langle \theta_0 f_1 \rangle \theta_0$$

$$\theta_1 = \frac{f_1'}{\|f_1'\|} \qquad \|f_1'\| = \sqrt{\langle f_1' f_1' \rangle}$$

$$f_j' = f_j - \langle \theta_0 f_j \rangle \theta_0 - \langle \theta_1 f_j \rangle \theta_1 \cdots \langle \theta_{j-1} f_j \rangle \theta_j$$

$$\theta_j = f_j' / \|f_j'\|$$

Vectors

$$a_1, a_2, a_3, \ldots a_m$$

$$q_1 = \frac{a_1}{\|a_1\|} \qquad \|a_1\| = \sqrt{a_1^T a_1}$$

$$a_2' = a_2 - (q_1^T a_2) q_1 \qquad q_2 = \frac{a_2'}{\|a_2\|}$$

$$\vdots$$

$$a_j' = a_j - (q_1^T a_j) q_1 - \cdots (q_{j-1}^T a_j) q_j$$

$$q_j = \frac{a_j'}{\|a_j\|}$$

Example Given polynomial basis functions $1, x, x^2, \ldots x^n$ in the range $[-1, 1]$ and $w(x) = 1$, determine corresponding orthonormal basis functions

(I) $\theta_0 = \dfrac{f_0}{\|f_0\|} = \dfrac{1}{\sqrt{\int_{-1}^{1} f_0^2(x)\,dx}} = \dfrac{1}{\sqrt{\int_{-1}^{1} 1\,dx}} = \boxed{\dfrac{1}{\sqrt{2}}}$

$\|f_0\| = \sqrt{\langle f_0, f_0 \rangle} = \sqrt{\int_{-1}^{1} f_0^2(x)\,dx} = \sqrt{2}$

(II) $f_1' = f_1 - \langle \theta_0 f_1 \rangle \theta_0$

$= x - \langle 1/\sqrt{2}, x \rangle\, 1/\sqrt{2}$ → odd

$= x - \left(\int_{-1}^{1} \frac{1}{\sqrt{2}} x \, dt\right) 1/\sqrt{2}$

$= x$

$\theta_1 = \dfrac{f_1'}{\|f_1'\|} = \dfrac{x}{\sqrt{\int_{-1}^{1} x^2\,dx}} = \dfrac{x}{\sqrt{2/3}}$

$$\boxed{\theta_1 = \sqrt{\frac{3}{2}}\, x}$$

(III) $f_2' = f_2 - \langle \theta_0 f_2 \rangle \theta_0 - \langle \theta_1 f_2 \rangle \theta_1$

$= x^2 - \left(\int_{-1}^{1} \frac{1}{\sqrt{2}} x^2\right) \cdot \frac{1}{\sqrt{2}} - \left(\int_{-1}^{1} \sqrt{\frac{3}{2}}\, x^3\right) \sqrt{\frac{3}{2}}\, x$

$= x^2 - \frac{1}{3}$

$$\boxed{\theta_2 = \frac{x^2 - 1/3}{\sqrt{\int_{-1}^{1} (x^2 - 1/3)\,dx}} = \sqrt{\frac{5}{2}}\,(3x^2 - 1)}$$

$\vdots$ θ_n

The functions $\theta_0, \theta_1, \dots \theta_n$ are called the Legendre Polynomials

In its typical representation, the Legendre polynomials are only proportional to θ_i's because by convention they are normalized to a length other than 1

$$P_n(x) = \sqrt{\frac{2}{2n+1}}\,\theta_n(x)$$

Legendre polynomials

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_{n+1}(x) = \frac{2n+1}{n+1}\,x\,P_n(x) - \frac{n}{n+1}P_{n-1}(x)$$

$$\langle P_i, P_j \rangle = \begin{cases} 0 & i \neq j \\ 2/2n+1 & i = j \end{cases}$$

Three term relationship

$$P_0(x) = 1$$

$$\underline{\underline{P_{n+1}}}(x) = (x - b_{n+1})\,\underline{\underline{P_n}}(x) - C_{n+1}\underline{\underline{P_{n-1}}}(x)$$

$$b_{n+1} = \frac{\langle xP_n, P_n \rangle}{\langle P_n, P_n \rangle} \qquad n = 0, 1, 2, \dots$$

$$c_{n+1} = \frac{\langle xP_n, P_{n-1} \rangle}{\langle P_{n-1}, P_{n-1} \rangle} \qquad n = 1, 2, \dots$$

$$C_1 = 0$$

Least squares regression with orthogonal basis functions

(a) Discrete (x_i, y_i) $i = 1, 2, \ldots N$

$$\hat{y}_i = \sum_{j=0}^{m} a_j \phi_j(x_i)$$ if ϕ_j's are orthogonal

$$\Phi^T \Phi a = \Phi^T Y$$

vector

$$\begin{bmatrix} \Phi_0^T \Phi_0 & \Phi_0^T \phi_1 & \cdots & \Phi_0^T \phi_m \\ \Phi_1^T \Phi_0 & \Phi_1^T \Phi_1 & \cdots & \Phi_1^T \phi_m \\ & & \ddots & \Phi_m^T \phi_m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \Phi_0^T Y \\ \Phi_1^T Y \\ \vdots \\ \Phi_m^T Y \end{bmatrix}$$

Off diagonal terms will be zero

$$a_j = \frac{\Phi_j^T Y}{\Phi_j^T \Phi_j}$$

(b) Continuous Case

$f(x)$ that has to be approximated by $g(x)$

$$g(x) = \sum_{j=0}^{m} a_j \phi_j(x)$$

$$\Phi^T \Phi a = \Phi^T Y$$

$$\begin{bmatrix} \langle \phi_0 \phi_0 \rangle & \langle \phi_0 \phi_1 \rangle & \cdots & \langle \phi_0 \phi_m \rangle \\ \langle \phi_1 \phi_0 \rangle & \langle \phi_1 \phi_1 \rangle & \cdots & \langle \phi_1 \phi_m \rangle \\ \vdots & & & \\ \langle \phi_1 \phi_m \rangle & - & - & \langle \phi_m \phi_m \rangle \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \langle \phi_0 f \rangle \\ \langle \phi_1 f \rangle \\ \vdots \\ \langle \phi_m f \rangle \end{bmatrix}$$

Unknowns

$$a_j = \frac{\langle \phi_j f \rangle}{\langle \phi_j \phi_j \rangle} = \frac{\int_a^b \phi_j(x) f(x)\, dx}{\int_a^b \phi_j^2(x)\, dx}$$

Discrete case

Example.

x	0	2	4	8
y	1	0.7937	0.63	0.3968

$$\boxed{\hat{y}_i = a_0 + a_1 x_i}$$

Linear

$$\Phi = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}$$

$$y = a_0 + a_1 x$$

$$P_0(x) = 1 = [1\ 1\ 1\ 1]^T$$

$$P_1(x) = (x - b_1) P_0(x) \qquad c_1 = 0$$

$$b_1 = \frac{\langle x P_0, P_0 \rangle}{\langle P_0, P_0 \rangle} = \frac{\langle x, 1 \rangle}{\langle 1, 1 \rangle} = \frac{x^T P_0}{P_0^T P_0}$$

$$= \frac{14}{4} = 3.5$$

$$P_1(x) = (x - 3.5)$$

$$P_2(x) = (x - 4.7837)(x - 3.5) - 8.75$$

Solving the problem with orthogonal basis

$$1, \quad (x - 3.5) \qquad y = a_0 + a_1 (x - 3.5)$$

$$\Phi = \begin{bmatrix} 1 & -3.5 \\ 1 & -1.5 \\ 1 & 0.5 \\ 1 & 4.5 \end{bmatrix}$$

$$\Phi^T \Phi = \begin{bmatrix} 4 & 0 \\ 0 & 35 \end{bmatrix} \qquad \Phi^T y = \begin{bmatrix} 2.8205 \\ 2.590 \end{bmatrix}$$

$$\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \phi_0^T y / \phi_0^T \phi_0 \\ \phi_1^T y / \phi_1^T \phi_1 \end{bmatrix} = \begin{bmatrix} 0.7051 \\ -0.0740 \end{bmatrix}$$

$$\hat{y}_i = a_0 P_0(x) + a_1 P_1(x)$$

$$= 0.7051 + (-0.074)(x - 3.5)$$

$$\boxed{\hat{y}_i = 0.9641 - 0.0740\, x_i}$$

<u>Example</u> Approximate $f(x) = e^x$ by a second order Legendre polynomial in the range $x \in (-1, 1)$

$$f(x) = e^x \qquad x \in (-1, 1)$$

$$g(x) = \sum_{j=0}^{2} a_j \, \phi_j(x)$$

ϕ_j's are Legendre Polynomials

$$\phi_0 = P_0(x) = 1$$

$$\phi_1 = P_1(x) = x$$

$$\phi_2 = P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$\Phi^T \Phi \, a = \Phi^T y$$

$$\Rightarrow \begin{bmatrix} \langle \phi_0, \phi_0 \rangle & & \\ & \langle \phi_1, \phi_1 \rangle & \\ & & \langle \phi_2, \phi_2 \rangle \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \langle \phi_0 f \rangle \\ \langle \phi_1 f \rangle \\ \langle \phi_2 f \rangle \end{bmatrix}$$

$$a_0 = \frac{\langle \phi_0 f \rangle}{\langle \phi_0 \phi_0 \rangle} = \frac{\int_{-1}^{1} 1 . e^x \, dx}{\int_{-1}^{1} 1^2 \, dx} = \frac{1}{2}\left[e - \frac{1}{e}\right]$$

$$= 1.1752$$

$$a_1 = \frac{\langle \phi_1 f \rangle}{\langle \phi_1 \phi_1 \rangle} = \frac{\int_{-1}^{1} x . e^x \, dx}{2/3} = \frac{3}{e}$$

$$= 1.1036$$

$$\langle P_i \, P_i \rangle = 2/_{2n+1}$$

$$a_2 = \frac{\langle \phi_2 f \rangle}{\langle \phi_2 \phi_2 \rangle} = \frac{\int_{-1}^{1} \frac{(3x^2-1)}{2} e^x \, dx}{2/5}$$

$$= \frac{5}{2}\left(e - 7/e\right) = 0.3578$$

$$g(x) = a_0 P_0(x) + a_1 P_1(x) + a_2 P_2(x)$$

$$g(x) = 0.9963 + 1.1036 + 0.5367x^2$$