Least square Problem

- Suppose a polynomial

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_m x_i^m + \varepsilon_i \ (i = 1,2,\ldots,n)$$

- In matrix form

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^m \\ 1 & x_2 & x_2^2 & \ldots & x_2^m \\ 1 & x_3 & x_3^2 & \ldots & x_3^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^m \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix},$$

- It is a linear problem

$$\overrightarrow{y} = \mathbf{X}\overrightarrow{\beta} + \overrightarrow{\varepsilon}.$$

Note that X is not square but rectangular because n is not equal to n

Solve via Inverse

$$\widehat{\vec{\beta}} = (\mathbf{X}^\top\mathbf{X})^{-1}\,\mathbf{X}^\top\vec{y},$$

It means best parameters to fit (least square criteria), in other word we are looking for the condition

$$\arg\min_{\beta} ||y - X\beta||,$$

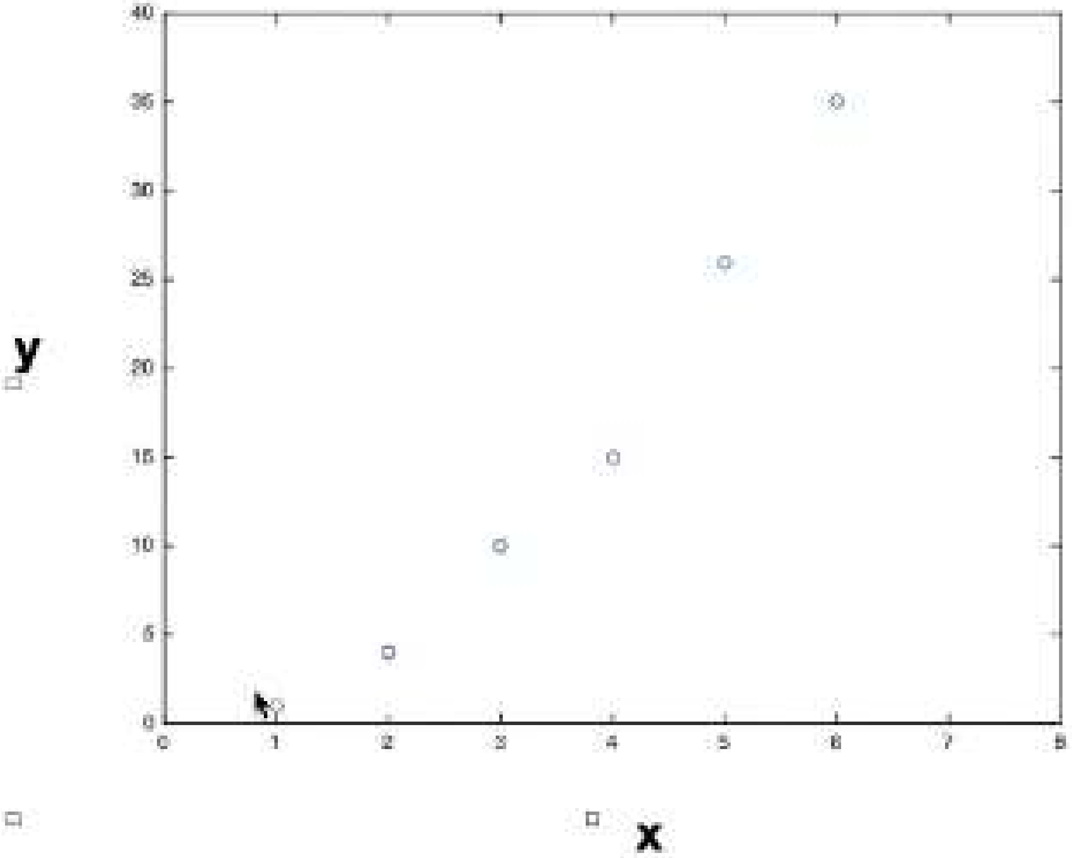
an **L² norm,** note that this norm will be zero for polynomial interpolation.

Lets evaluate

- Calculate parameter β_0, β_1 and β_2 (Regression problem)

Data

x	y
1	1
2	4
3	10
4	15
5	26
6	35



Lets evaluate

- Assembly the rectangular (not square) Vandermonde Matrix

Data

x	y
1	1
2	4
3	10
4	15
5	26
6	35

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix}$$

nxm =6x3, *n* **data** *m* **unknown**

First

- Assembly the rectangular (not square) Vandermonde Matrix

$$\widehat{\overrightarrow{\beta}} = (\mathbf{X}^{\top}\mathbf{X})^{-1}\,\mathbf{X}^{\top}\overrightarrow{y},$$

M

$$\mathbf{X}^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 9 & 16 & 25 & 36 \end{bmatrix} \qquad \mathbf{X} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} 6 & 21 & 91 \\ 21 & 91 & 441 \\ 91 & 441 & 2275 \end{bmatrix}$$ *a 3x3 need to find the inverse*

Lets evaluate Parameter

$$\widehat{\vec{\beta}} = (\underbrace{\mathbf{X}^{\mathsf{T}}\mathbf{X}}_{\mathbf{M}})^{-1}\,\mathbf{X}^{\mathsf{T}}\vec{y},$$

$$\mathbf{M} = \begin{bmatrix} 6 & 21 & 91 \\ 21 & 91 & 441 \\ 91 & 441 & 2275 \end{bmatrix}$$

a *3x3* need to find the inverse $\mathbf{M}^{-1}$

Using the LU decomposition above, resulted

$$\mathbf{L} = \begin{bmatrix} 6 & 0 & 0 \\ 21 & 17.5 & 0 \\ 91 & 122.5 & 37.33 \end{bmatrix} \qquad \mathbf{U} = \begin{bmatrix} 1 & 3.5 & 15.17 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{bmatrix}$$

Upper Matrix

$$\mathbf{M}\mathbf{M}^{-1} = \mathbf{I}$$

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} m'_{11} & m'_{12} & m'_{13} \\ m'_{21} & m'_{22} & m'_{23} \\ m'_{31} & m'_{32} & m'_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{L}\underbrace{\mathbf{U}\mathbf{M}^{-1}}_{\mathbf{D}} = \mathbf{I}$$

$$\underbrace{\begin{bmatrix} 6 & 0 & 0 \\ 21 & 17.5 & 0 \\ 91 & 122.5 & 37.33 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} 1 & 3.5 & 15.17 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m'_{11} & m'_{12} & m'_{13} \\ m'_{21} & m'_{22} & m'_{23} \\ m'_{31} & m'_{32} & m'_{33} \end{bmatrix}}_{\mathbf{D}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{I}}$$

Lower Matrix

$$\mathbf{MM}^{-1} = \mathbf{I}$$

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} m'_{11} & m'_{12} & m'_{13} \\ m'_{21} & m'_{22} & m'_{23} \\ m'_{31} & m'_{32} & m'_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{LD} = \mathbf{I}$$

$$\begin{bmatrix} 6 & 0 & 0 \\ 21 & 17.5 & 0 \\ 91 & 122.5 & 37.33 \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

D-Matrix

$$\begin{bmatrix} 6 & 0 & 0 \\ 21 & 17.5 & 0 \\ 91 & 122.5 & 37.33 \end{bmatrix} \begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- By forward substitution

$$\begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{bmatrix} = \begin{bmatrix} 0.1667 & 0 & 0 \\ -0.2000 & 0.0571 & 0 \\ 0.2500 & -0.1875 & 0.0268 \end{bmatrix}$$

The Inverse

$$\mathbf{UM}^{-1} = \mathbf{D}$$

$$\begin{bmatrix} 1 & 3.5 & 15.17 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m'_{11} & m'_{12} & m'_{13} \\ m'_{21} & m'_{22} & m'_{23} \\ m'_{31} & m'_{32} & m'_{33} \end{bmatrix} = \begin{bmatrix} 0.1667 & 0 & 0 \\ -0.2000 & 0.0571 & 0 \\ 0.2500 & -0.1875 & 0.0268 \end{bmatrix}$$

- By backward substitution

$$\begin{bmatrix} m'_{11} & m'_{12} & m'_{13} \\ m'_{21} & m'_{22} & m'_{23} \\ m'_{31} & m'_{32} & m'_{33} \end{bmatrix} = \begin{bmatrix} 3.2 & -1.95 & 0.2500 \\ -1.95 & 1.3696 & -0.1875 \\ 0.25 & -0.1875 & 0.0268 \end{bmatrix}$$

Best Parameters

$$\mathbf{M}^{-1} = \begin{bmatrix} m'_{11} & m'_{12} & m'_{13} \\ m'_{21} & m'_{22} & m'_{23} \\ m'_{31} & m'_{32} & m'_{33} \end{bmatrix} = \begin{bmatrix} 3.2 & -1.95 & 0.2500 \\ -1.95 & 1.3696 & -0.1875 \\ 0.25 & -0.1875 & 0.0268 \end{bmatrix}$$

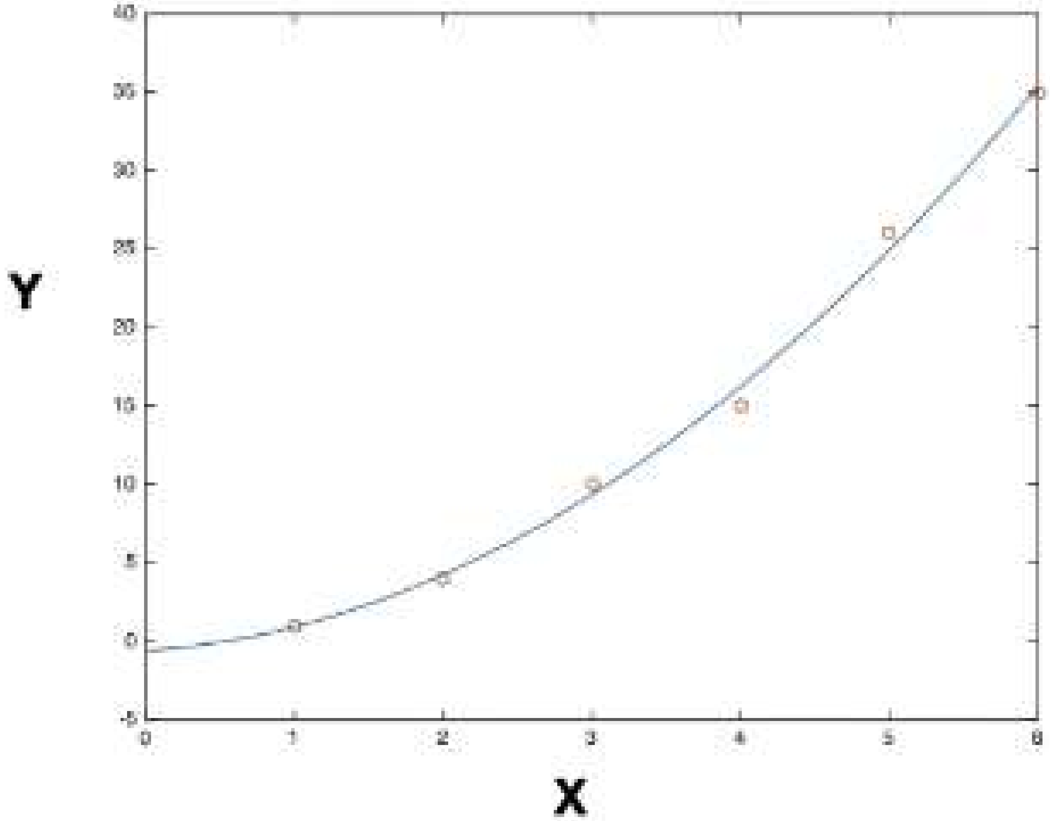
$$\widehat{\overrightarrow{\beta}} = \begin{bmatrix} 3.2 & -1.95 & 0.2500 \\ -1.95 & 1.3696 & -0.1875 \\ 0.25 & -0.1875 & 0.0268 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 9 & 16 & 25 & 36 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 10 \\ 15 \\ 26 \\ 35 \end{bmatrix}$$

$$\widehat{\overrightarrow{\beta}} = \begin{bmatrix} -0.6 \\ 0.6357 \\ 0.8929 \end{bmatrix}$$

Plug in into polynomial

$$\widehat{\vec{\beta}} = \begin{bmatrix} -0.6 \\ 0.6357 \\ 0.8929 \end{bmatrix}$$

$$y_i = -0.6 + 0.6357 x_i + 0.8929 x_i^2$$



Python — Vim crout.py — 100×25

```python
import numpy as np

def crout(A):
    n = A.shape[0]

    U = np.zeros((n, n), dtype=np.double)
    L = np.zeros((n, n), dtype=np.double)

    for k in range(n):

        L[k, k] = A[k, k] - L[k, :] @ U[:, k]

        U[k, k:] = (A[k, k:] - L[k, :k] @ U[:k, k:]) / L[k, k]
        L[(k+1):, k] = (A[(k+1):, k] - L[(k+1):, :] @ U[:, k]) / U[k, k]

    return L, U

A = np.array([[1, 4, 5], [6, 8, 22], [32, 5., 5]])
L,U=crout(A)
print(A)
print(L)
print(U)
print(L@U
~
"crout.py" 23L, 494B                                    23,10          All
```