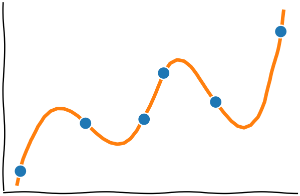
Interpolation and approximation

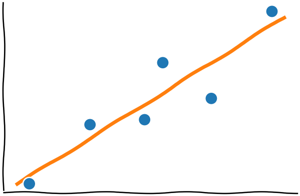
Given a set of points $\{(x_j, y_j), j = 1, \cdots, n\}$, and given a functional form $f(x; \vec{\beta})$, find *"best"* $\vec{\beta}$ so that $f(x; \vec{\beta})$ *"models"* the data.

Interpolation



$$f(x_j; \vec{\beta}) = y_j$$

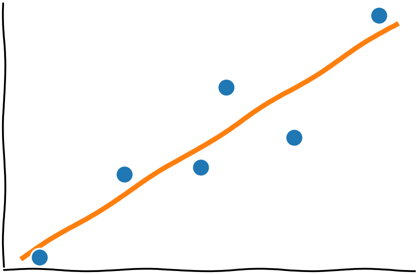
Approximation



$$f(x_j; \vec{\beta}) + \varepsilon_j = y_j$$

ε_j is "noise", $\mathbb{E}(\varepsilon_j) = 0$

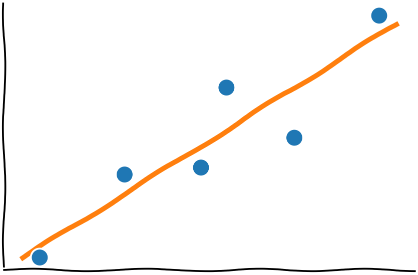
Least squares approximation



$$\mathbb{E}(\varepsilon_j) = 0$$
$$\mathbb{E}(\varepsilon_j^2) = \sigma_j^2$$
$$\sigma_j \approx \text{const}$$

$$\chi^2(\vec{\beta}) = \sum_{j=1}^{n} \left| y_j - f(x_j; \vec{\beta}) \right|^2 \Rightarrow \min$$

Weighted least squares

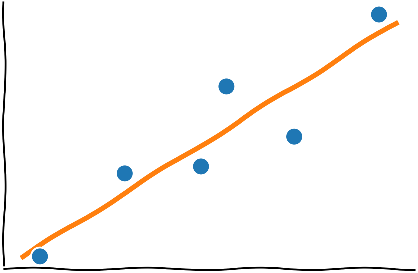


$$\mathbb{E}(\varepsilon_j) = 0$$
$$\mathbb{E}(\varepsilon_j^2) = \sigma_j^2$$

If σ_j are significantly different (*heteroscedasticity*)

$$\chi^2(\vec{\beta}) = \sum_{j=1}^{n} \frac{\left| y_j - f(x_j; \vec{\beta}) \right|^2}{\sigma_j^2} \Rightarrow \min$$

Least absolute deviations

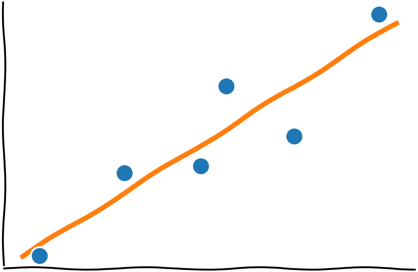


Also known as L_1 regression:

$$S(\vec{\beta}) = \sum_{j=1}^{n} \left| y_j - f(x_j; \vec{\beta}) \right| \Rightarrow \min$$

Total least squares

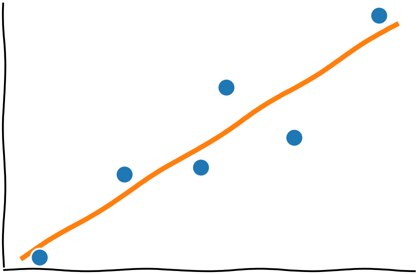
Also known as *Orthogonal distance regression*: minimize the sum of squares of orthogonal distances from observations to the curve.



Can be more appropriate *e.g.* if both variables, x and y have measurement errors.

Linear least squares

Least squares approximation



$$\mathbb{E}(\varepsilon_j) = 0$$
$$\mathbb{E}(\varepsilon_j^2) = \sigma_j^2$$
$$\sigma_j \approx \text{const}$$

$$\chi^2(\vec{\beta}) = \sum_{j=1}^{n} \left| y_j - f(x_j; \vec{\beta}) \right|^2 \Rightarrow \min$$

Linear least squares

Consider an ordinary least squares problem,

$$\xi(\vec{\beta}) = \sum_{j=1}^{n} \left| y_j - f(x_j; \vec{\beta}) \right|^2 \Rightarrow \min$$

Let the model, $f(x;\beta)$, is a *linear* function of $\vec{\beta}$, a linear combination of m *basis functions*, $\varphi_k(x)$

$$f(x; \vec{\beta}) = \sum_{j=1}^{m} \beta_k \varphi_k(x)$$

Typically, want $m < n$.

Linear least squares

Let the model, $f(x; \beta)$, is a *linear* function of $\vec{\beta}$, a linear combination of m *basis functions*, $\varphi_k(x)$

$$f(x; \vec{\beta}) = \sum_{j=1}^{m} \beta_k \varphi_k(x)$$

The basis functions need not be linear:

- polynomials: $\varphi_k(x) = x^k$
- Fourier series: $\varphi_k(x) = e^{is_k x}$
- $\varphi_k(x) = x^k \log x$
- $\ldots$

Linear least squares

We minimize with respect to $\vec{\beta}$

$$\xi(\vec{\beta}) = \sum_{j=1}^{n} |z_j|^2$$

where ($j = 1, \ldots, n$)

$$z_j = y_j - \left(\beta_1 \varphi_1(x_j) + \beta_2 \varphi_2(x_j) + \cdots + \beta_m \varphi_m(x_j)\right)$$

Which is eqivalent to

$$\xi(\beta) = \left\|\mathbf{y} - \mathbf{A}\vec{\beta}\right\|_2^2$$

with $\mathbf{y} = (y_1, \cdots, y_n)^T$ and $A_{kj} = \varphi_k(x_j)$.

Design matrix

The *design matrix* $\mathbf{A}$ is an $n \times m$ matrix

$$A = \begin{bmatrix} \varphi_1(\quad) & \varphi_2(\quad) & \cdots & \varphi_m(\quad) \\ \varphi_1(\quad) & \varphi_2(\quad) & \cdots & \varphi_m(\quad) \\ & & \cdots & \\ \varphi_1(\quad) & \varphi_2(\quad) & \cdots & \varphi_m(\quad) \end{bmatrix}$$

The dimensions of the design matrix is *# of observations* $\times$ *# of parameters*

Design matrix

The *design matrix* $\mathbf{A}$ is an $n \times m$ matrix

$$A = \begin{bmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_m(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_m(x_2) \\ & & \cdots & \\ \varphi_1(x_n) & \varphi_2(x_n) & \cdots & \varphi_m(x_n) \end{bmatrix}$$

The dimensions of the design matrix is *# of observations* $\times$ *# of parameters*

Example: straight line fit

The model is

$$f(x;\vec{\beta}) = \beta_1 + \beta_2\, x$$

$$m = 2: \qquad \varphi_1(x) = 1\,, \qquad \varphi_2(x) = x$$

and the design matrix is

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

Linear least squares

Normal equations

Linear least squares: normal equations

To minimize the quadratic form

$$\xi(\vec{\beta}) = \left\|\mathbf{y} - \mathbf{A}\vec{\beta}\right\|_2^2$$

set the derivatives to zero,

$$\frac{\partial}{\partial \beta_k}\xi(\vec{\beta}) = 0\,, \qquad\qquad j = 1, \cdots, m$$

And obtain the *normal equations*:

$$\mathbf{A}^T \mathbf{A}\,\vec{\beta} = \mathbf{A}^T \mathbf{y}$$

Linear least squares: normal equations

Normal equations

$$\mathbf{A}^T \mathbf{A}\, \vec{\beta} = \mathbf{A}^T \mathbf{y}$$

give a formal solution of a linear least squares problem.

However,

$$\mathrm{cond}\left(\mathbf{A}^T \mathbf{A}\right) = [\mathrm{cond}\, A]^2$$

so that typically the system of normal equations is *very* poorly conditioned.

Linear least squares

QR factorization of the design matrix

Linear least squares: QR factorization

Recall that a matrix $\mathbf{A}$ can be factorized into

$$\mathbf{A} = \mathbf{Q}\mathbf{R}$$

where $\mathbf{Q}$ is orthogonal ($\mathbf{Q}^T\mathbf{Q} = \mathbf{1}$) and $\mathbf{R}$ is upper triangular.

Since a design matrix is thin and tall ($m < n$), last $n - m$ rows of $\mathbf{R}$ are zero:

$$\mathbf{A} = \mathbf{Q}\begin{bmatrix}\mathbf{R_1} \\ \mathbf{0}\end{bmatrix}$$

where $\dim \mathbf{R}_1 = m$

Linear least squares: QR factorization

Since the 2-norm of a vector is invariant under a rotation by an orthogonal matrix $\mathbf{Q}$, we rotate the residual $\mathbf{y} - \mathbf{A}\vec{\beta}$

$$\xi(\beta) = \left\| \mathbf{y} - \mathbf{A}\vec{\beta} \right\|^2 = \left\| \mathbf{Q}^T \left(\mathbf{y} - \mathbf{A}\vec{\beta} \right) \right\|^2$$

$$= \left\| \mathbf{Q}^T \mathbf{y} - \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} \vec{\beta} \right\|^2$$

Next, write

$$\mathbf{Q}^T \mathbf{y} = \begin{bmatrix} \mathbf{f} \\ \mathbf{r} \end{bmatrix}$$

with $\dim \mathbf{f} = m$ and $\dim \mathbf{r} = n - m$.

Linear least squares: QR factorization

This way,

$$\xi(\vec{\beta}) = \left\|\mathbf{f} - \mathbf{R}_1\vec{\beta}\right\|^2 + \|\mathbf{r}\|^2$$

And the minimum of $\xi(\vec{\beta})$ satisfies

$$\mathbf{R}_1\vec{\beta} = \mathbf{f}$$

Linear least squares: QR factorization

This way,

$$\xi(\vec{\beta}) = \left\|\mathbf{f} - \mathbf{R}_1\vec{\beta}\right\|^2 + \|\mathbf{r}\|^2$$

And the minimum of $\xi(\vec{\beta})$ satisfies

$$\mathbf{R}_1\vec{\beta} = \mathbf{f}$$

Algorithm

- Factorize the design matrix $\mathbf{A} = \mathbf{QR}$
- Rotate $\mathbf{y} \to \mathbf{Q}^T\mathbf{y}$ (only need m rows $\Rightarrow$ thin QR)
- Solve $\mathbf{R}_1\vec{\beta} = \mathbf{f}$ by back substitution.