Homework 10

Least Squares.

The meaning of linear.

Least squares approximation is also called *Linear Regression*. The meaning of *linear* here is not that the modeling fuction, say $f(x)$, must be linear, or rather affine (linear + shift), but that the solution is a linear combination of the modeling *basis functions*:

$$f(x) = \beta_0 f_0(x) + \beta_1 f_1(x) + \cdots + \beta_n f_n(x)$$

Exercise 1. Fitting with a Parabola.

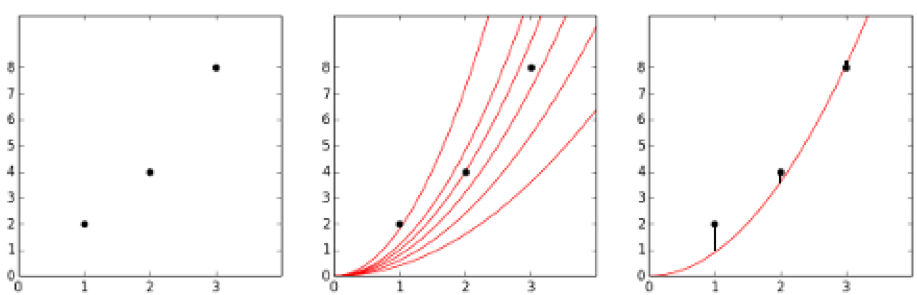
Here is a set of data points (x_i, y_i) that we want to model with a parabola.

$$S = \{(1, 2), (2, 4), (3, 8)\}$$

This is a one dimensional model, there is no shift. The parabola will go through the origin and the modeling function will be a one dimensional "linear combination" of the shape:

$$f(x) = \beta_0 x^2$$

Which one is the best fitting parabola? Which one minimizes $\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 = |\varepsilon\|^2$?



We sample the model function, x^2, above the independent variable x which gives us the unique column of the design matrix $X = (1, 4, 9)^T$. So we are now solving $X\beta_0 + \varepsilon = y$ in order to minimize the square of the norm of ε.

$$\begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} \beta_0 + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}$$

To find the best β_0 we take the orthogonal projection of y on the column space of X.

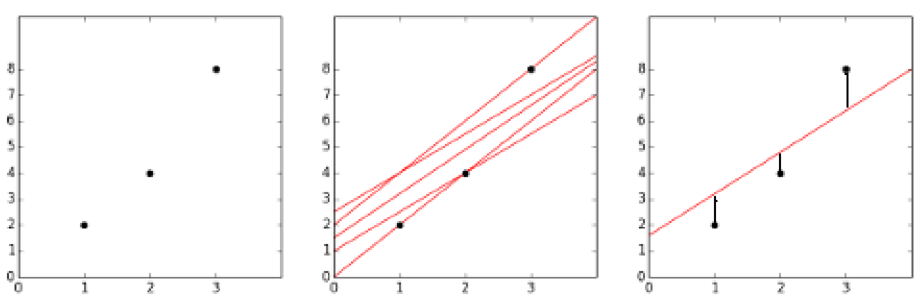
i) Recall the formula giving the orthogonal projection matrix P_X.

ii) From $X\hat{\beta} = P_X y$, find the optimal parameter $\hat{\beta}_0$ minimizing $\|\varepsilon\|^2$.

Exercise 2. Fitting with a straight Line.

Here is the classical exemple of least squares approximation where the regression is linear, we have a linear combination of the basis functions $f_0(x) = 1$ and $f_1(x) = x$ and the model is affine, a straight line through the cloud of data points, with a shift at the origin.

$$f(x) = \beta_0 + \beta_1 x$$



Let us take the data points from Exercise 1., $S = \{(1, 2), (2, 4), (3, 8)\}$.

i) Compute the orthogonal projection matrix P_X either by hand or with Matlab/Julia/R (or even better first by hand, then checking with a computer).

You can use the following formula to invert a 2×2 matrix by hand.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

ii) Compute $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)^T$.