

An Exploration of ‘Polar’ Regression

A better label for this topic is desired...

David Fleischer

Last Update: 13 February, 2018

Introduction

(Ordinary) Linear Regression

We recall that in linear regression we assume our responses $Y \in \mathbb{R}^n$ are *linearly* related to our $\mathbb{X} \in \mathbb{R}^{n \times p}$, i.e.,

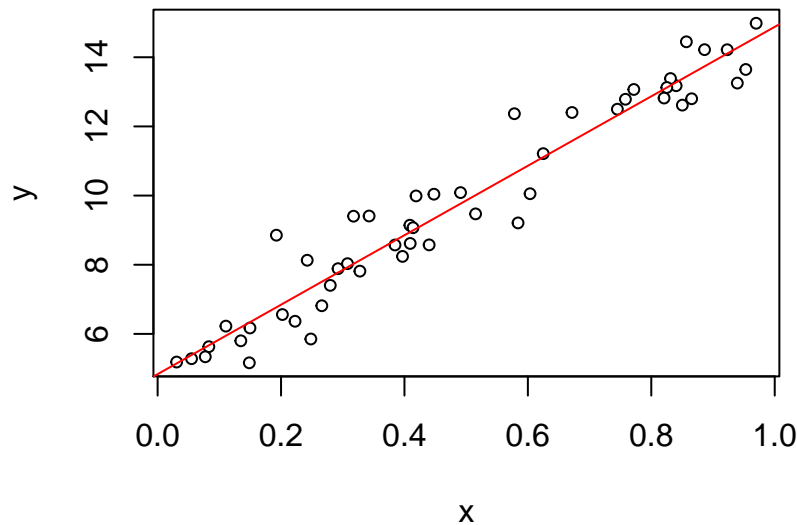
$$Y = \mathbb{X}\beta + \varepsilon,$$

such that $\beta \in \mathbb{R}^p$ and ε is some random perturbation with zero mean. Indeed, the typical problem of linear regression is the problem of solving for the ordinary least squares estimator of β , denoted by $\hat{\beta}^{\text{OLS}}$, given by

$$\hat{\beta}^{\text{OLS}} = \arg \min_{\beta} \{ \|Y - \mathbb{X}\beta\|_2^2 \}.$$

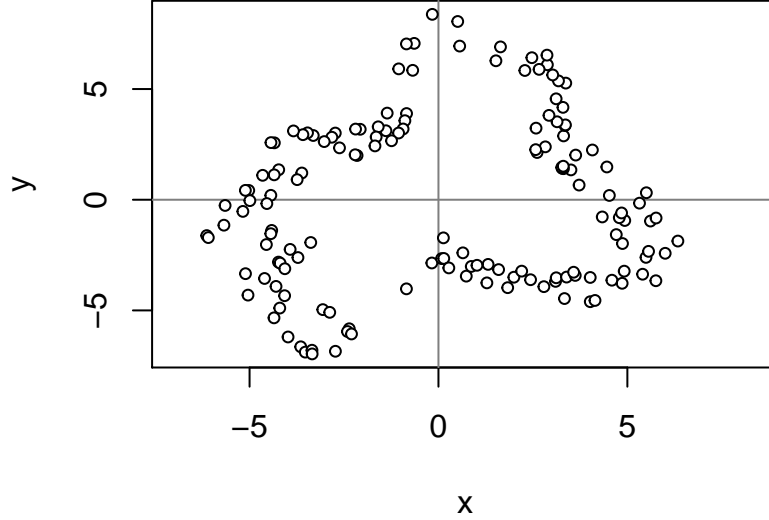
In the univariate case we typically visualize such a problem as the problem of finding a line of best fit across pairs $\{(x_i, y_i)\}_{i=1}^n$. As a quick visual,

Ordinary Linear Fit

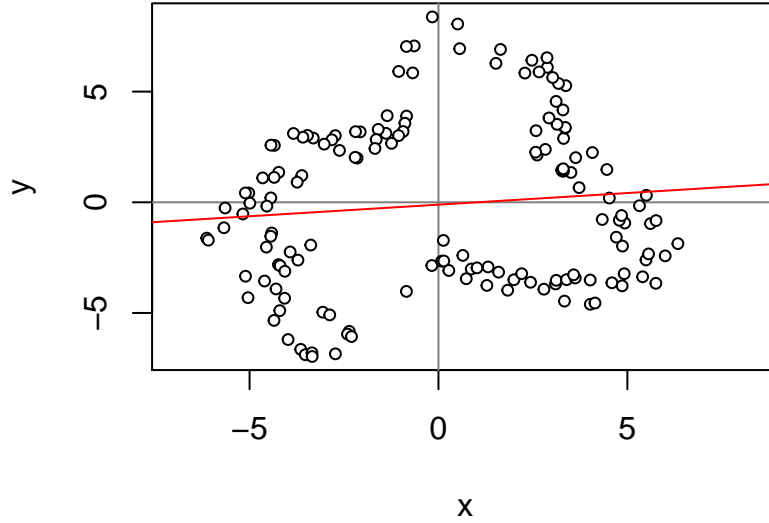


An Illustrative Example

Consider the problem of finding a fit for the following points:



Clearly such a problem is ill-suited for the tools of linear regression. Computing a linear fit on our points $\{(x_i, y_i)\}_{i=1}^n$ generates the disastrous fit



Despite not being a *linear* relationship, it's clear to the practitioner that some radial relationship exists as we traverse through a rotation about the origin.

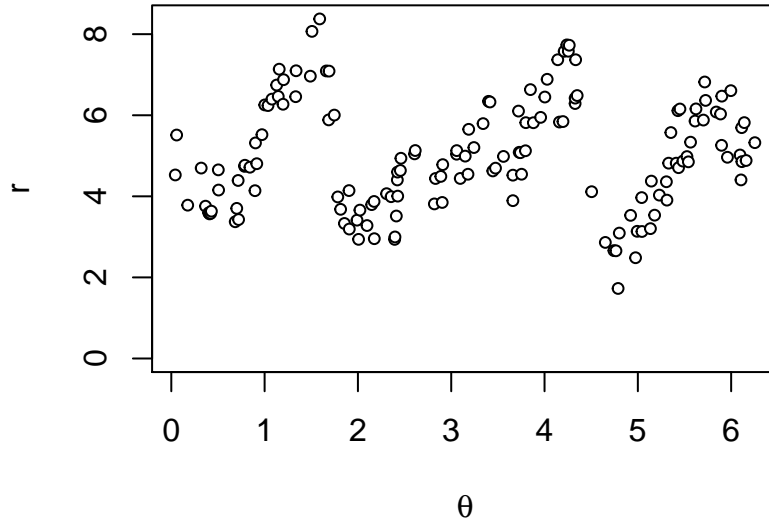
Attacking the Problem through Polar Coordinates

An immediate problem highlighted by the above example is that our linear model attempts to fit a linear function to data that almost certainly cannot describe a function. One remedy to this problem is to consider the Cartesian-to-polar transformation

$$r = \sqrt{x^2 + y^2}$$

$$\theta = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } x > 0, y > 0, \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if } x < 0, y > 0, \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if } x < 0, y < 0, \\ \arctan\left(\frac{y}{x}\right) + 2\pi & \text{if } x > 0, y < 0. \end{cases}$$

Under such a coordinate transformation may “unravel” our curve from Cartesian coordinates into points in polar coordinates. Using the above example we find our data to be better behaved (in the sense that the mean is now described by *some* function $r(\theta)$)



In this coordinate system, our noisy data can be modelled by a relationship of the form

$$r = f(\theta) + \varepsilon,$$

subject to $f(0) = f(2\pi)$,

where the additional constraint $f(0) = f(2\pi)$ ensures that our curve in polar coordinates describes a closed curve in Cartesian coordinates. In our running example, our function $r = r(\theta)$ was given by

$$r(\theta) = 5 + \theta(\theta - 2\pi)(0.23 \sin(3\theta) + 0.21 \sin(4\theta) + 0.07 \sin(7\theta) + 0.05 \sin(11\theta)) + \varepsilon$$

for $\varepsilon \sim \mathcal{U}(-1, 1)$, whose mean function $\mathbb{E}[r | \theta]$ is plotted below:

