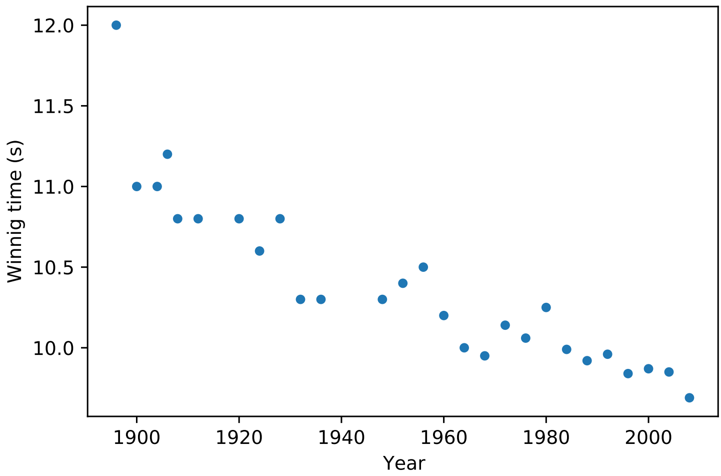
Polynomial regression and basis functions

Herman Kamper

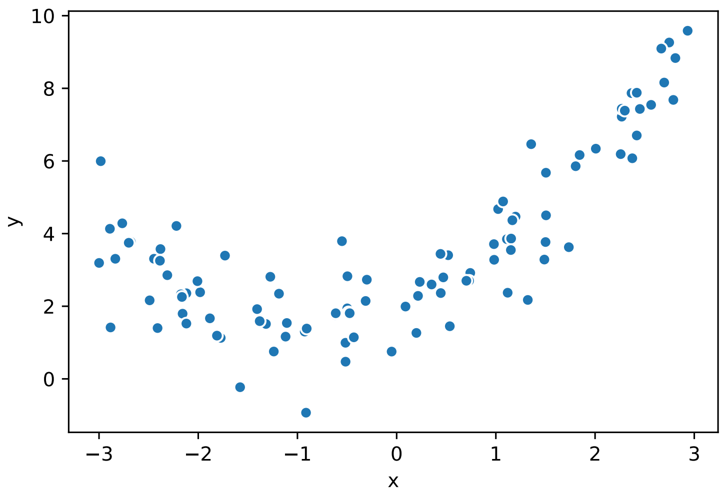
2024-01, CC BY-SA 4.0

Linear and non-linear relationships

Linear relationship:



Non-linear relationship:



From multiple linear regression to polynomial regression

Multiple linear regression recap

Model:

$$f(\mathbf{x}; \mathbf{w}) = w_0 + w_1 x_1 + w_2 x_2 + \ldots + w_D x_D = \mathbf{w}^\top \mathbf{x}$$

Fit on data $\left\{(\mathbf{x}^{(n)}, y^{(n)})\right\}_{n=1}^{N}$ using the squared loss:

$$\begin{aligned} J(\mathbf{w}) &= \sum_{n=1}^{N} \left(y^{(n)} - f(\mathbf{x}^{(n)}; \mathbf{w})\right)^2 \\ &= (\mathbf{y} - \mathbf{X}\mathbf{w})^\top (\mathbf{y} - \mathbf{X}\mathbf{w}) \end{aligned}$$

with

$$\mathbf{X} = \begin{bmatrix} — \left(\mathbf{x}^{(1)}\right)^\top — \\ — \left(\mathbf{x}^{(2)}\right)^\top — \\ \vdots \\ — \left(\mathbf{x}^{(N)}\right)^\top — \end{bmatrix} \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(N)} \end{bmatrix}$$

Solution:

$$\hat{\mathbf{w}} = \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{y}$$

Polynomial regression

What if we want to fit:

$$f(x; w_0, w_1, w_2) = w_0 + w_1 x + w_2 x^2$$

Let us define:

$$\boldsymbol{\phi}(x) =$$

We can then write:

$$f(x; \mathbf{w}) =$$

Now we can solve this problem exactly as for multiple linear regression by pretending that $\boldsymbol{\phi}(\mathbf{x})$ is $\mathbf{x}$.

Our design matrix now becomes:

$$\boldsymbol{\Phi} =$$

Example with multivariate input

For

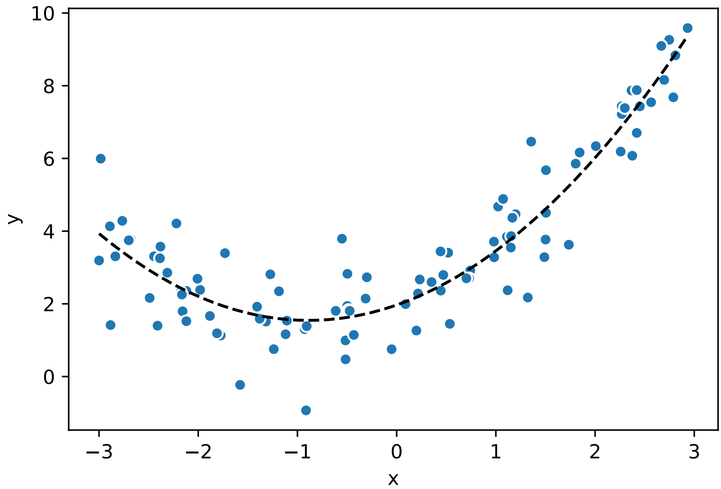
$$f(\mathbf{x}; \mathbf{w}) = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1 x_2 + w_4 x_1^2 + w_5 x_2^2$$

we would have

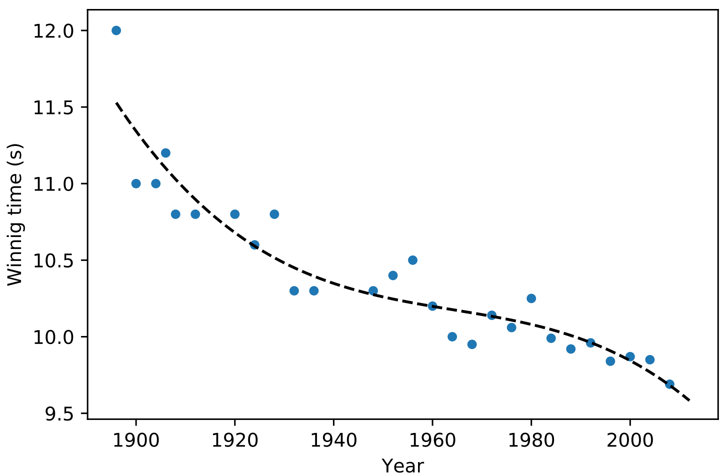
$$\boldsymbol{\phi}(\mathbf{x}) =$$

Polynomial regression examples

Quadratic polynomial:



Third-order polynomial:



Basis functions

Instead of just polynomials, we can actually put any function in

$$\boldsymbol{\phi}(\mathbf{x}) = \begin{bmatrix}\phi_1(\mathbf{x}) & \phi_1(\mathbf{x}) & \dots & \phi_K(\mathbf{x})\end{bmatrix}^\top$$

E.g. $\sin$, $\cos$, $\log$, $\exp$, FFT, etc.

This can be quite useful if we have some inside domain knowledge of the data and the problem.

Radial basis function (RBF)

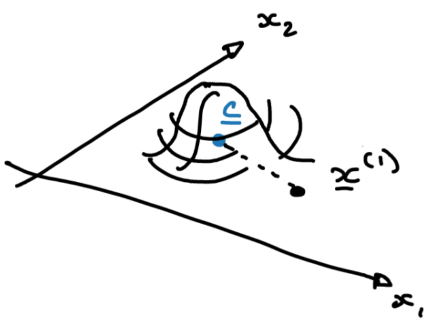
An RBF is a function whose value depends on the distance between the input and some fixed point:

$$\phi(\mathbf{x}) = \exp\left\{\frac{-(\mathbf{x}-\mathbf{c})^\top(\mathbf{x}-\mathbf{c})}{h^2}\right\}$$

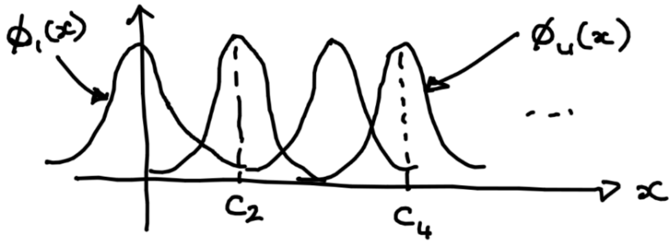
Intuitively, it measures how far input $\mathbf{x}$ is from $\mathbf{c}$, with h controlling how much we penalise points that are far way.

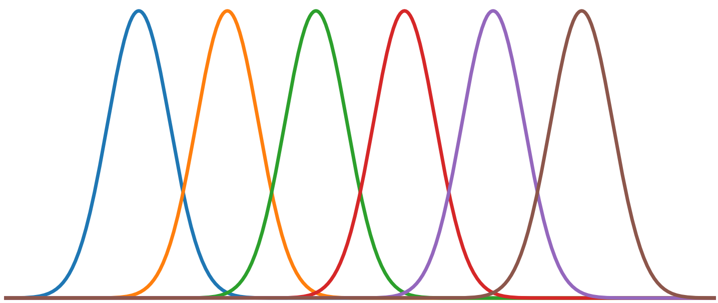
RBF in one dimension:

RBF in two dimensions:



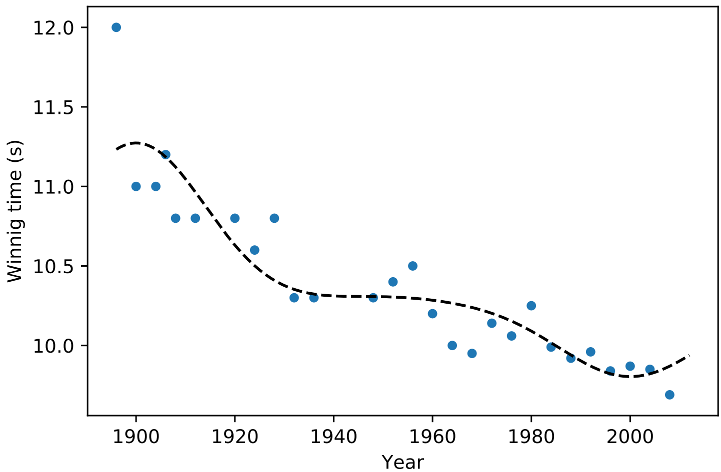
Can even have a family of RBFs:



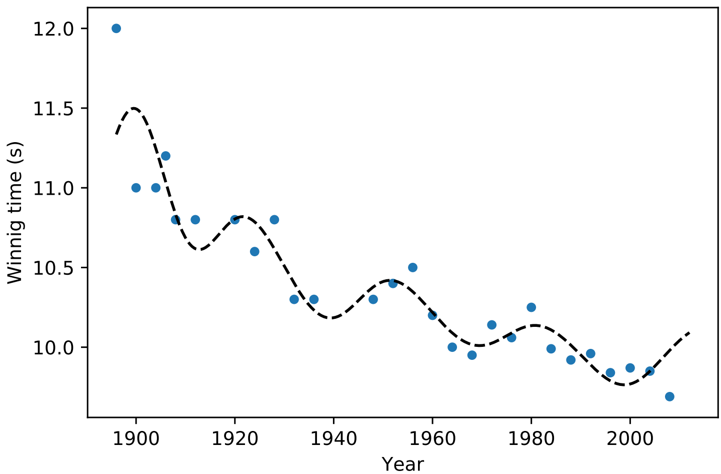


RBF basis functions examples

RBF with $c = [1900, 1950, 2000]$ and $h = 20$:



RBF with $c = [1900, 1910, \ldots, 2000]$ and $h = 10$:



Videos covered in this note

- Linear regression 3: Polynomial regression and basis functions (15 min)

Reading

- ISLR 3.3.2
- ISLR 7.1
- ISLR 7.3

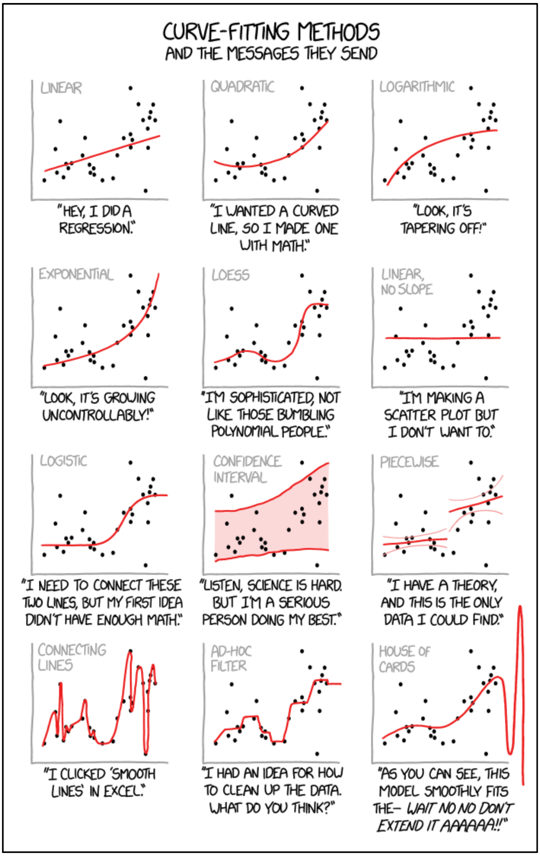


Figure from https://xkcd.com/2048/.