**Splines and Degrees of Freedom**

Splines are piecewise polynomials which all for additional flexibility in basis functions. They work by creating relationships between low order polynomials and stitching them together.

More complexity is added to the spline prediction of the true relationship as the degrees of freedom increases.

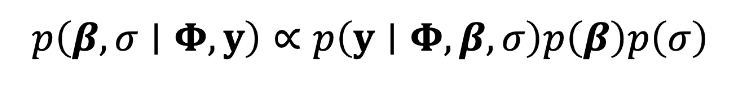
At this point, we will be assuming an unknown sigma. Our goal is to pick the best most and determine the relationship between the RMSE value (sigma) of each model and the added complexity as the degrees of freedom increases.

A flexible model will be capable of achieving a lower RMSE than a simple **biased** model.

**Bayesian Linear Basis Function Models**

Goal: Learn posterior distribution on the unknown -parameters and the unknown noise ()

* -parameters are a-priori independent of the noise



* Due to the conditional Independence, the likelihood function can be factored
* Likelihood functions for a linear model are **Gaussian**!

Logo

Description automatically generated with medium confidence

* For the nth observation, the mean is

Common Prior Specifications:

Common to assume all coefficients are independent a-priori.

Diagram

Description automatically generatedShared prior mean and shared prior standard deviation

🡪 “no prior bias”

Complete Bayesian Linear Basis Function

Likelihood:

Deterministic Relationship:

Text

Description automatically generated

Diagram

Description automatically generatedPrior on Regression Coefficients:

Prior on :

Text

Description automatically generated with medium confidence

Helps determine if RMSE is going to be high or low in posterior.

Another common practice is to use a separate prior standard deviation for the intercept ()

* Intercept given its own set of hyperparameters:
* A picture containing text, furniture, table

  Description automatically generatedBasis feature slopes, , use common hyperparameters:
* Prior standard deviation on intercept often larger than prior standard deviation on regression coefficients.
  + Intercept has wider range of values that it can be

Unknown Values in Basis Functions

* Degrees of Freedom: Number of features deriving
* Depends on the degree of the function
  + For basis function of degree J, have J+1 unknown coefficients (intercept + each feature)
  + Total J + 2 unknowns 🡪 J + 1 unknown coefficients + sigma (RMSE of the model)
* Columns/Features of Basis Design Matrix
  + Each spline “feature” represents a polynomial relationship over the range of the input x
  + Each feature is focused on a different area of the entire input range
* For a 4 df spline

Text

Description automatically generated with medium confidence

* + -parameters
    - act as weights to each of the polynomials (features)
    - Greater the absolute value of the weight 🡪 polynomial has greater impact on the mean trend (Weighted contribution)
    - are CONSTANTS 🡪 DO NOT DEPEND ON INPUT
      * Linearly related to the mean
* As Degree of Freedom increases, each feature has a narrower focus over a small range of the input

Laplace Approximation

* Need to apply a transformation on to use Laplace
  + Must ensure bounds of are not violated
* Because the model is linear, the initial guess does not matter
  + ONLY ONE UNIQUE MAXIMUM
* As degrees of freedom increases posterior distribution on decreases (getting pushed towards zero)
  + Can arbitrarily make model as good as possible