x

Posterior

From Wikipedia (https://en.wikipedia.org/wiki/Multivariate_normal_distribution#Conditional_distributions)

is linear in the observations .

- We have a linear predictor! In fact it is the best linear unbiased predictor (BLUP).

  is lower than the marginal variance .

- SÅenaßa tQe bßi`

- Data can be noisy.
- The amplitude of the function is not 2 (i.e., we don't know it in advance).
- Correlation doesn't decay uniformly in all directions (i.e., radially).
- Even the most ideally smooth physical relationships are rarely *infinitely smooth*.

GP hyperparameters

Scale

Lets suppose you wanted your prior to generate random functions which had an amplitude larger than two.

- You could introduce a scale parameter

How, then, do we estimate the hyperparameter α?

So actually we have where the product is component-wise, and contains a matrix of Euclidean distances.

The full derivative

So we have

```
gradnl <- function(par, D, Y)
 {
    theta <- par[1]; g
```

g