Hierarchical model construction

1. Classic hierarchical Bayesian models
2. Gaussian process models
3. Projects

Occupancy models

$$z_i \sim Bernoulli(\psi_i)$$

$$y_{ij} \sim Bernoulli(z_i p)$$

Sites $i, ..., N$
Repeat visits $j, ..., J$

N-mixture models

$$N_i \sim Poisson(\lambda_i)$$

$$y_{ij} \sim Binomial(N_i, p)$$

Error in variables models

$$y_i \sim Normal(\alpha + \beta \tilde{x}_i, \sigma_y)$$

$$x_i \sim Normal(\tilde{x}_i, \sigma_x)$$

Example: modeling poop at ponds

$$y_i \sim Poisson(\mu_i)$$

$$\mu_i = \alpha_0 + log(\tilde{\pi})$$

$$\pi \sim N(\tilde{\pi}, \sigma_\pi)$$

Zero inflated Poisson

$$p(y_i|\theta, \lambda) = \begin{cases} \theta + (1-\theta) Poisson(0 \mid \lambda) & \text{if } y = 0 \\ (1-\theta) Poisson(y_i \mid \lambda) & \text{if } y > 0 \end{cases}$$

θ: mixing parameter

Zero inflated gamma

$$p(y_i|\theta, \lambda) = \begin{cases} \theta & \text{if } y = 0 \\ (1-\theta)\, Gamma(y_i \mid \alpha, \beta) & \text{if } y > 0 \end{cases}$$

θ: mixing parameter

Beta glm

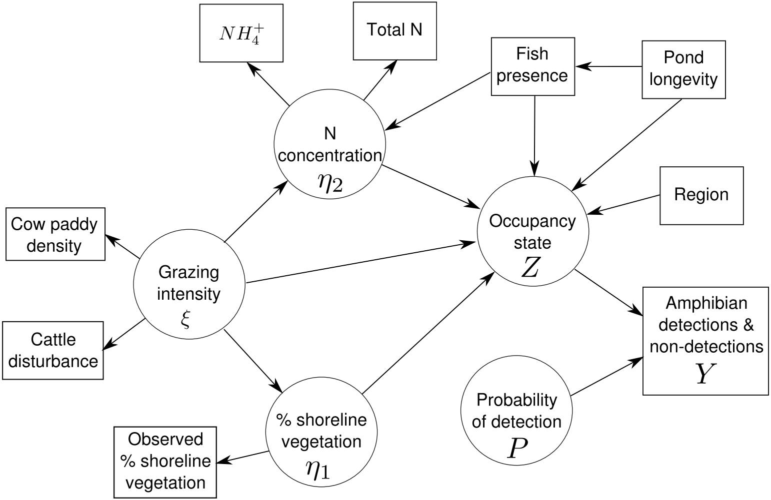
$$y_i \sim Beta(\alpha, \beta)$$

$$\alpha = \mu\phi$$

$$\beta = (1 - \mu)\phi$$

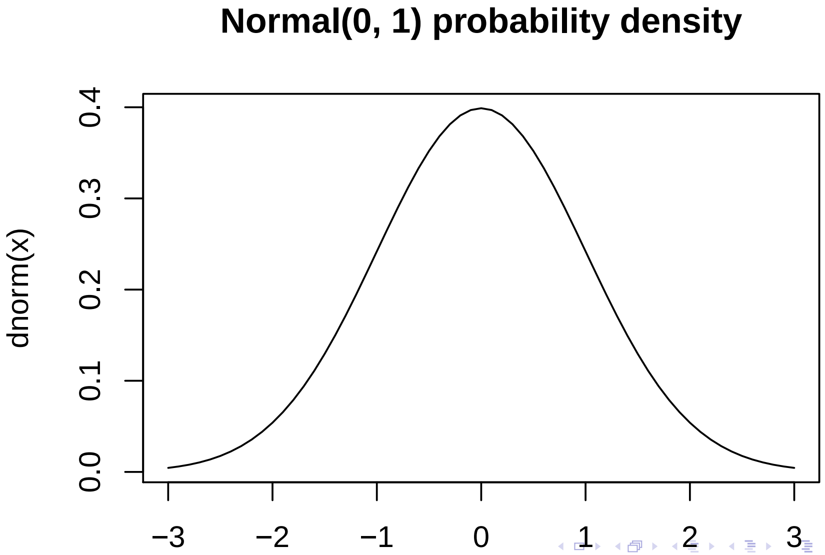
$$logit(\mu) = X\beta$$

Hierarchical Bayesian structural equation models



Background: univariate normal

$$x \sim N(\mu, \sigma^2)$$



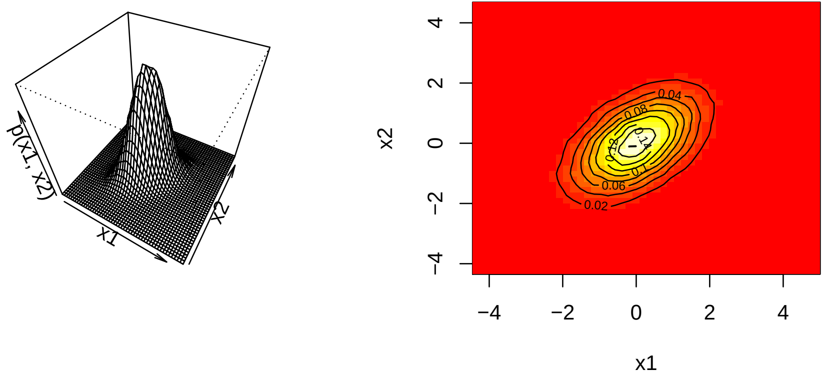
Background: multivariate normal

$$\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$\boldsymbol{\mu}$: vector of means
$\boldsymbol{\Sigma}$: covariance matrix

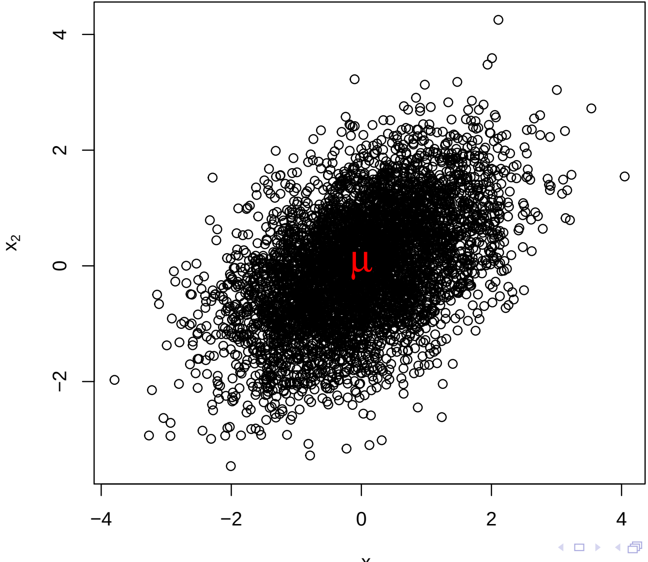
Bivariate normal probability density

$$\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$



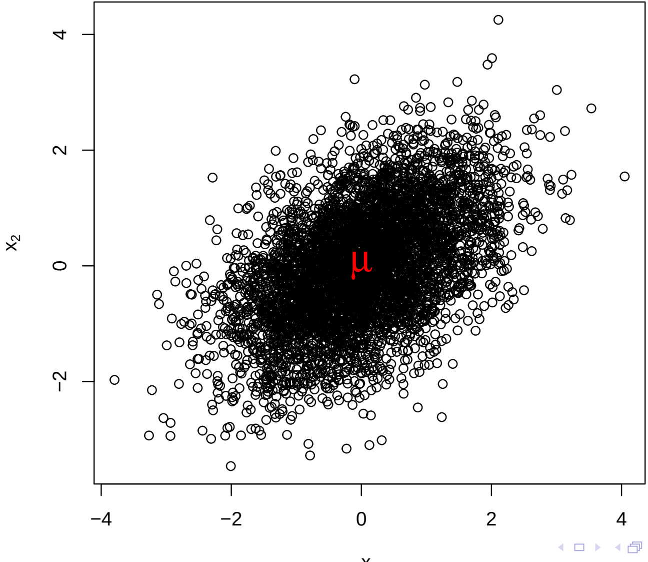
Bivariate normal parameters

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} Cov[X_1, X_1] & Cov[X_1, X_2] \\ Cov[X_2, X_1] & Cov[X_2, X_2] \end{bmatrix}$$



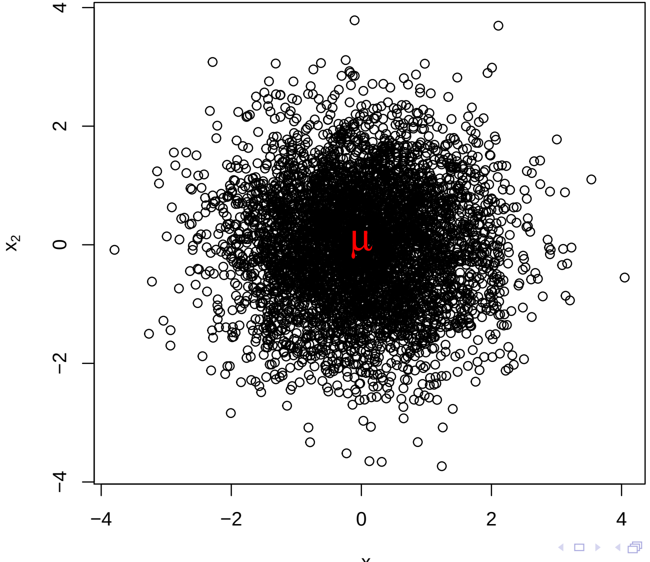
Bivariate normal parameters

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \; \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$



Uncorrelated bivariate normal

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



Common notation

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \; \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

Common notation

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \ \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$Cov[X_1, X_1] = Var[X_1] = \sigma_1^2$

$Cov[X_1, X_2] = \rho\sigma_1\sigma_2$

Σ must be symmetric and positive semi-definite

Classic linear modeling

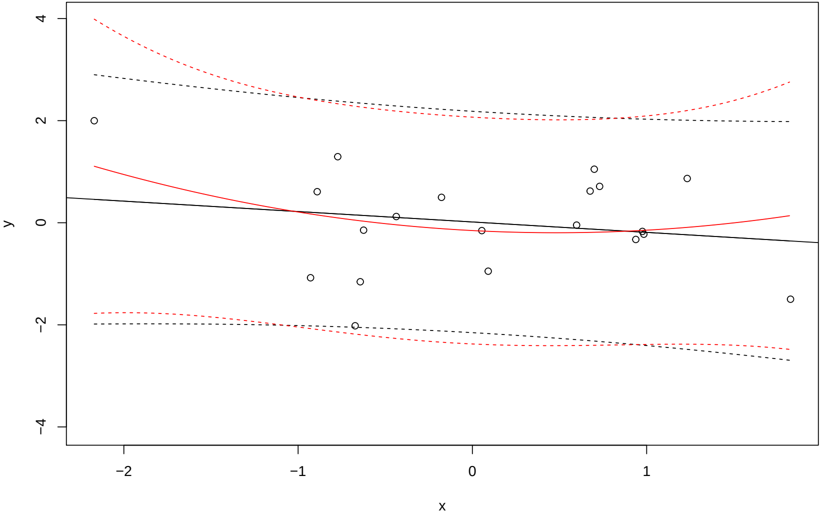
$$y = X\beta + \epsilon$$

$$\epsilon \sim N(0, \sigma^2)$$

Functional form determined by $X\beta$

Linear model functional forms

e.g. $y = \mu(x) + \epsilon$



Why not set a prior on $\mu(x)$?

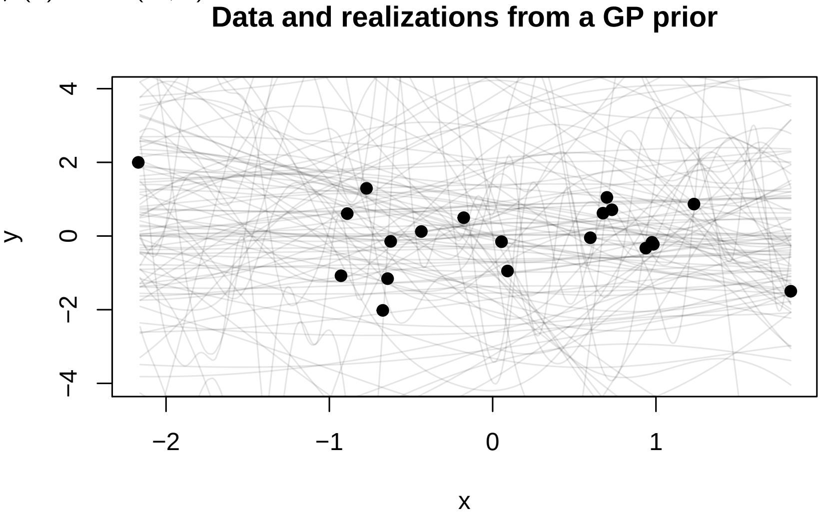
Gaussian process as a prior for $\mu(x)$

$$y \sim N(\mu(x), \sigma^2)$$

$$\mu(x) \sim GP(m, k)$$

GP prior for $\mu(x)$

$$y \sim N(\mu(x), \sigma^2)$$
$$\mu(x) \sim GP(m, k)$$



Wait, what's Gaussian about that?

If $\mu(x) \sim GP(m, k)$, then
$\mu(x_1), ..., \mu(x_n) \sim N(m(x_1), ..., m(x_n), K(x_1, ..., x_n)$
m and k are functions!

Mean function: m

Classic example: $m(x) = X\beta$
e.g., $\mu(x) \sim GP(X\beta, k(x))$
But, the covariance function $k(x)$ is the real star.

Covariance functions

k specifies covariance between to x values
Squared exponential covariance:

$$k(x, x') = \tau^2 exp\Big(-\frac{|x - x'|^2}{\phi}\Big)$$

Lots of options: smooth, jaggety, periodic

Example of squared exponential

$$\boldsymbol{K} = \begin{bmatrix} \tau^2 exp(-\frac{|x_1 - x_1|^2}{\phi}) & \tau^2 exp(-\frac{|x_1 - x_2|^2}{\phi}) \\ \tau^2 exp(-\frac{|x_2 - x_1|^2}{\phi}) & \tau^2 exp(-\frac{|x_2 - x_2|^2}{\phi}) \end{bmatrix}$$

Example of squared exponential

$$\boldsymbol{K} = \begin{bmatrix} \tau^2 exp(-\frac{0^2}{\phi}) & \tau^2 exp(-\frac{|x_1 - x_2|^2}{\phi}) \\ \tau^2 exp(-\frac{|x_2 - x_1|^2}{\phi}) & \tau^2 exp(-\frac{0^2}{\phi}) \end{bmatrix}$$

Example of squared exponential

$$\boldsymbol{K} = \begin{bmatrix} \tau^2 exp(0) & \tau^2 exp(-\frac{|x_1 - x_2|^2}{\phi}) \\ \tau^2 exp(-\frac{|x_2 - x_1|^2}{\phi}) & \tau^2 exp(0) \end{bmatrix}$$

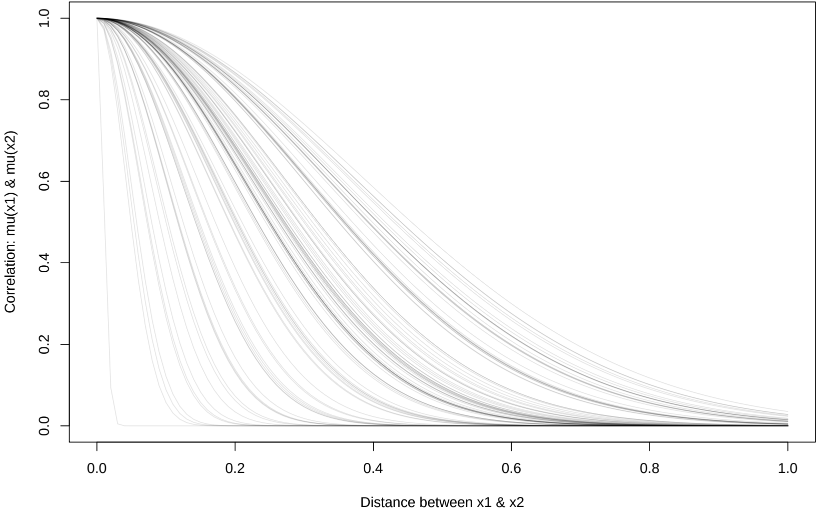
Example of squared exponential

$$\boldsymbol{K} = \begin{bmatrix} \tau^2 & \tau^2 exp(-\frac{|x_1 - x_2|^2}{\phi}) \\ \tau^2 exp(-\frac{|x_2 - x_1|^2}{\phi}) & \tau^2 \end{bmatrix}$$

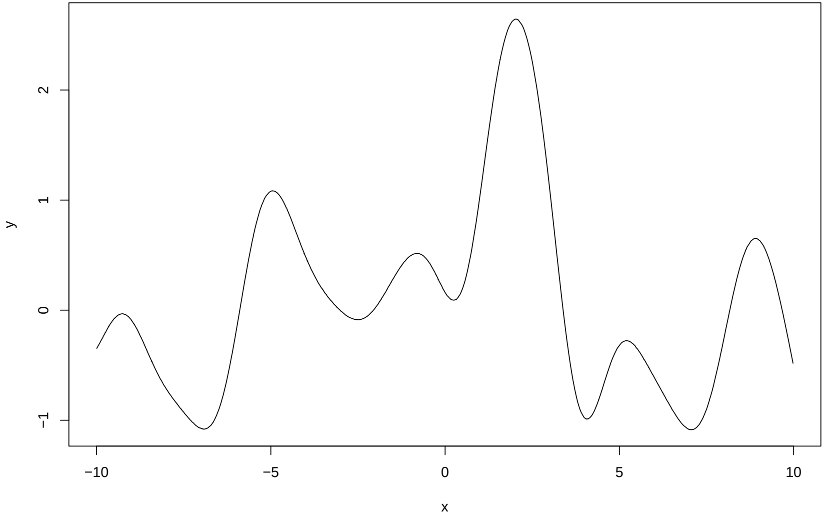
$Cor(\mu(x_1), \mu(x_2)) = exp(-\frac{|x_1 - x_2|^2}{\phi})$.

Correlation function

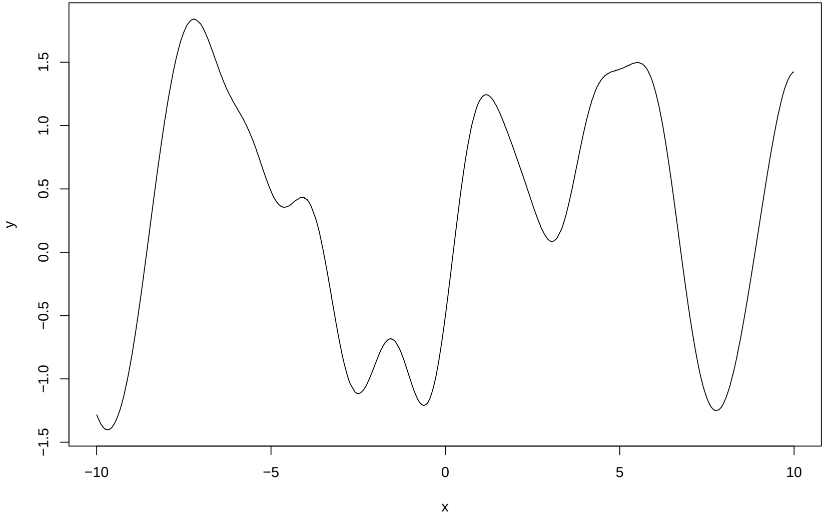
$$Cor(\mu(x_1), \mu(x_2)) = exp(-\frac{|x_1 - x_2|^2}{\phi}).$$



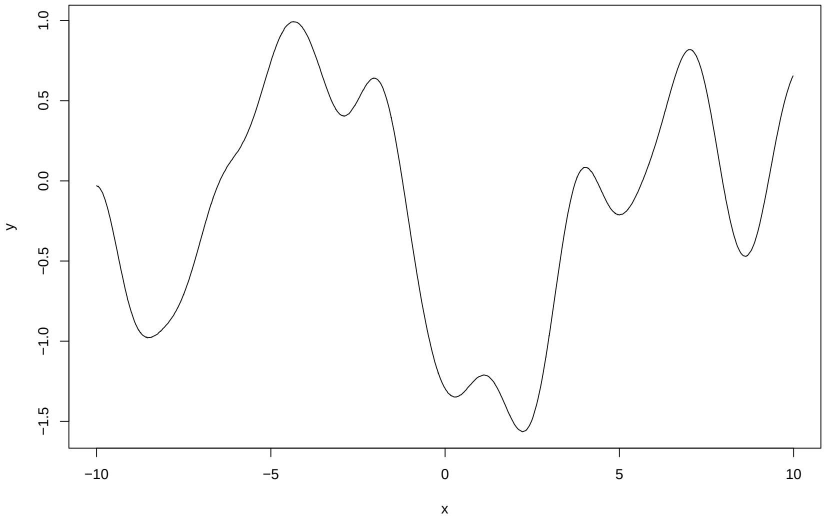
Gaussian process realizations



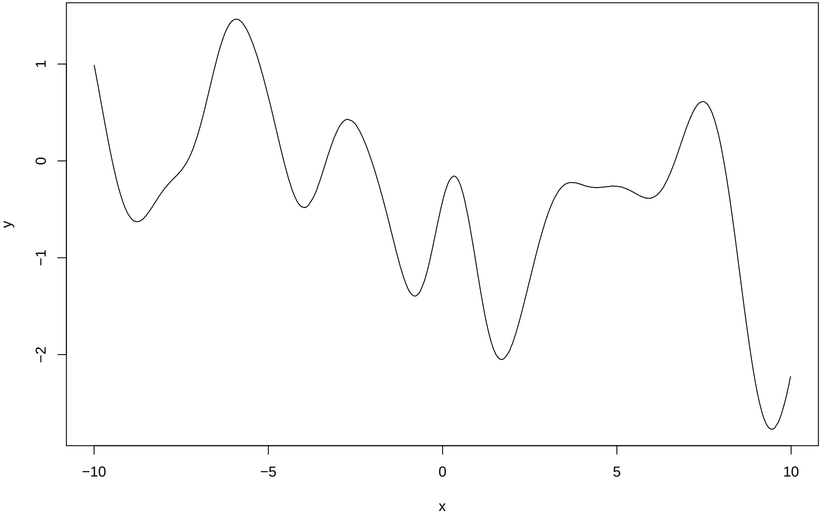
Gaussian process realizations



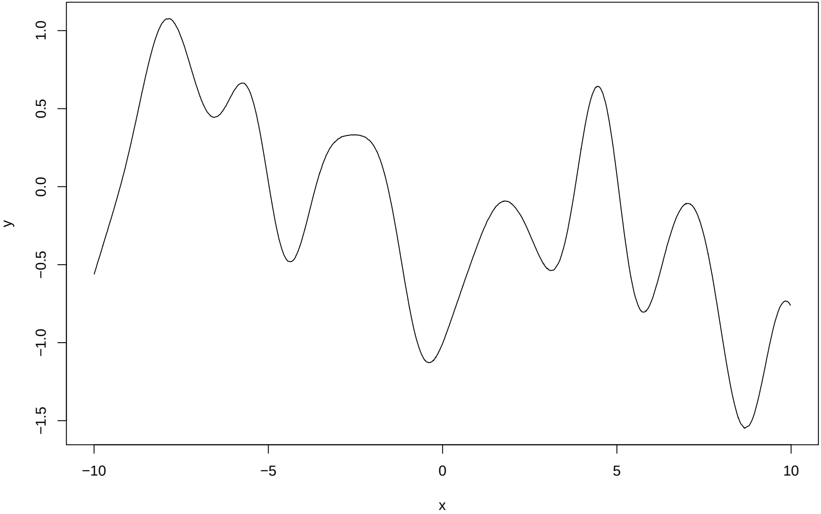
Gaussian process realizations



Gaussian process realizations

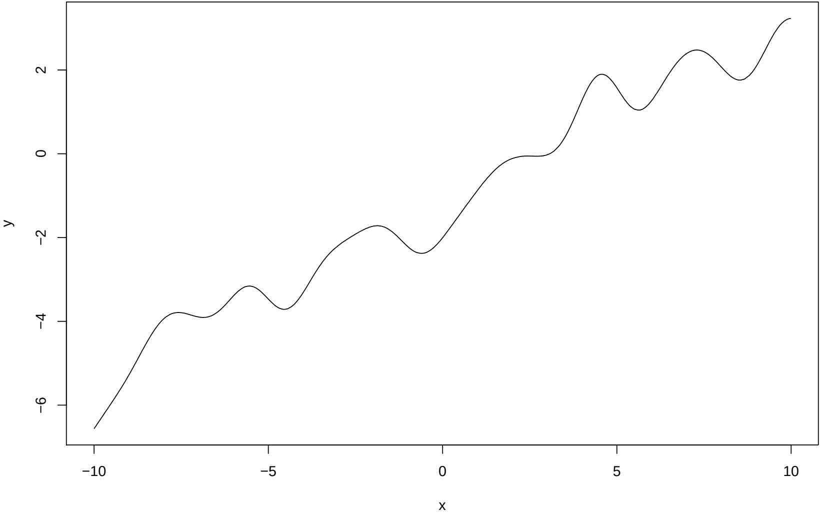


Gaussian process realizations



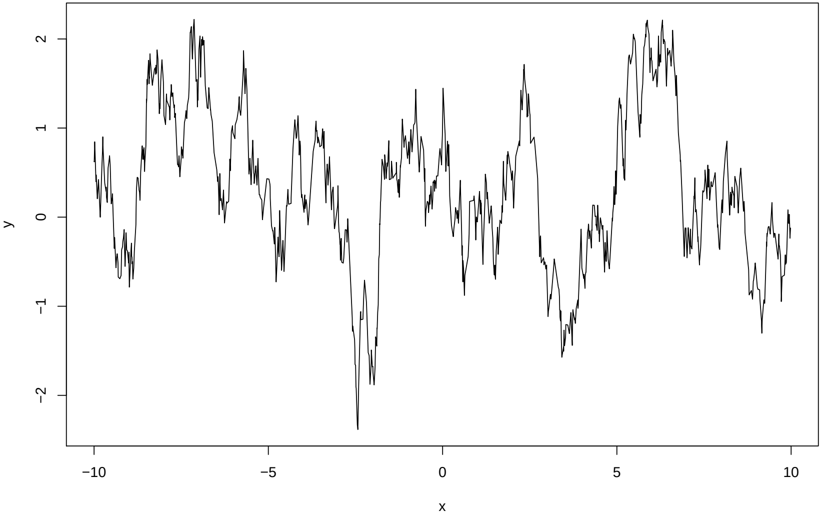
Gaussian process with nonzero mean function

$y \sim N(\mu, \sigma_y)$
$\mu \sim GP(X\beta, k(x))$



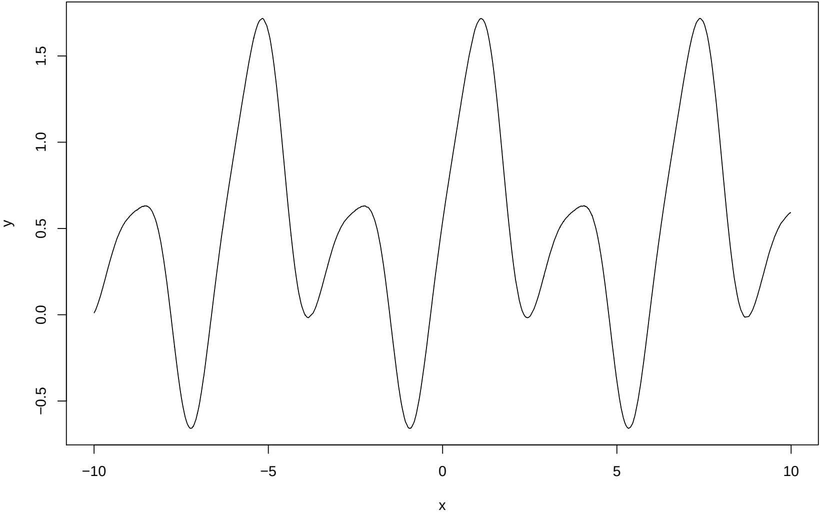
Ornstein–Uhlenbeck Gaussian process

$$k(x_1, x_2) = e^{\frac{d_{x_1,x_2}}{\phi}}$$



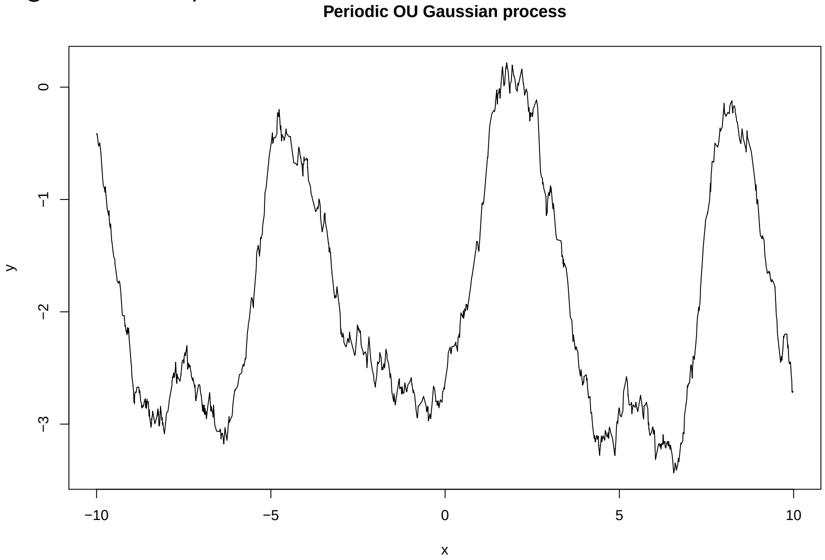
Periodic Gaussian process

$$k(x_1, x_2) = exp(\frac{2\sin^2(d/2)}{\phi})$$

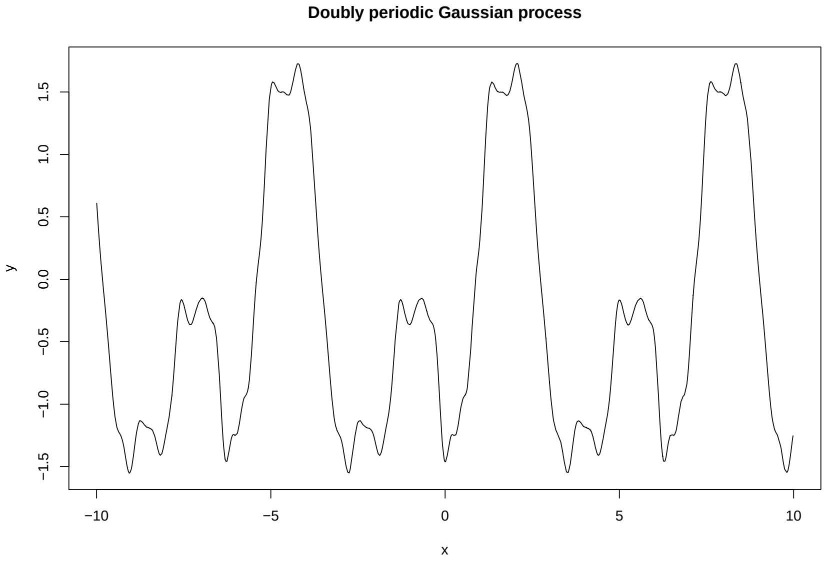


Combining Gaussian processes

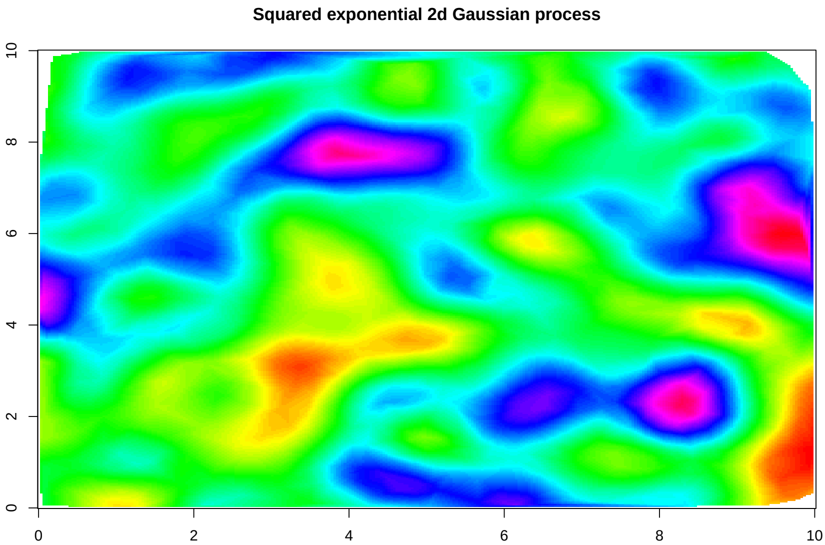
e.g., sums and products of covariance functions



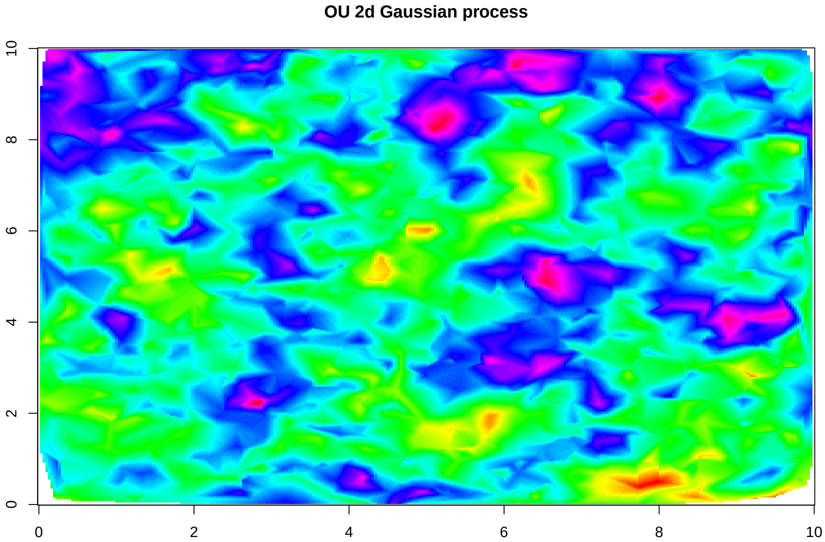
Combining Gaussian processes



Multidimensional inputs



Multidimensional inputs



Other inputs

Generally, $k(x)$ maps **distance** to **correlation**

- phylogenetic distance $\rightarrow$ phylogenetic correlation
- pedigree distance $\rightarrow$ additive genetic correlation
- distance in time $\rightarrow$ temporal correlation

Student projects

Before fitting your model to your data:

1. Write out model in mathematical notation (ideally LaTeX)
2. Prior predictive simulations (do your priors make sense?)
3. Model verification (given known parameters from PPS, do you recover parameters?)