Chapter 20: Basis function models
Chapter 21: Gaussian process models
Discussion of homework due beginning of Class 11b

- Computing problem
- Computing problem
- Applied problem
Discussion of homework due beginning of Class 11b

- Computing problem
- Computing problem
- Applied problem
Discussion of homework due beginning of Class 11b

- Computing problem
- Computing problem
- Applied problem
Discussion of homework due beginning of Class 11b

- Computing problem
- Computing problem
- Applied problem
Computing problem

- Probit regression
- Variational Bayes using latent-data formulation
Computing problem

- Probit regression
- Variational Bayes using latent-data formulation
Computing problem

- Probit regression
- Variational Bayes using latent-data formulation
Computing problem

- Well-switching in Bangladesh
- HMC or Metropolis for probit model
Computing problem

- Well-switching in Bangladesh
- HMC or Metropolis for probit model
Computing problem

- Well-switching in Bangladesh
- HMC or Metropolis for probit model
Applied problem

- Hierarchical model for voting by age and state
- Also education, also income
Hierarchical model for voting by age and state
Also education, also income
Applied problem

- Hierarchical model for voting by age and state
- Also education, also income
20. Basis function models

- Splines and basis functions
- Discrete or continuous models
- Multivariate regression surfaces
20. Basis function models

- Splines and basis functions
- Discrete or continuous models
- Multivariate regression surfaces
20. Basis function models

- Splines and basis functions
- Discrete or continuous models
- Multivariate regression surfaces
20. Basis function models

- Splines and basis functions
- Discrete or continuous models
- Multivariate regression surfaces
20.1. Splines and weighted sums of basis functions

- Cubic spline basis functions
- Gaussian kernels
- Nonnormal errors
20.1. Splines and weighted sums of basis functions

- Cubic spline basis functions
- Gaussian kernels
- Nonnormal errors
20.1. Splines and weighted sums of basis functions

- Cubic spline basis functions
- Gaussian kernels
- Nonnormal errors
20.1. Splines and weighted sums of basis functions

- Cubic spline basis functions
- Gaussian kernels
- Nonnormal errors
21 cubic B-splines
Random draws from spline model with independent normal $N(0,1)$ priors for coefs
Linear and spline fits to data
20.2. Basis selection and shrinkage of coefficients

- Similar to prior distributions for regression coefficients
- Scale mixtures of normals
- Behavior near 0 and tail behavior
20.2. Basis selection and shrinkage of coefficients

- Similar to prior distributions for regression coefficients
- Scale mixtures of normals
- Behavior near 0 and tail behavior
20.2. Basis selection and shrinkage of coefficients

- Similar to prior distributions for regression coefficients
- Scale mixtures of normals
- Behavior near 0 and tail behavior
20.2. Basis selection and shrinkage of coefficients

- Similar to prior distributions for regression coefficients
- Scale mixtures of normals
- Behavior near 0 and tail behavior
20.3. Non-normal models and multivariate regression surfaces

- Mixture models for errors
- Additive models
- Multivariate kernels
- Tensor products
20.3. Non-normal models and multivariate regression surfaces

- Mixture models for errors
- Additive models
- Multivariate kernels
- Tensor products
20.3. Non-normal models and multivariate regression surfaces

- Mixture models for errors
- Additive models
  - Multivariate kernels
- Tensor products
20.3. Non-normal models and multivariate regression surfaces

- Mixture models for errors
- Additive models
- Multivariate kernels
- Tensor products
20.3. Non-normal models and multivariate regression surfaces

- Mixture models for errors
- Additive models
- Multivariate kernels
- Tensor products
Sum of discretized prediction ‘tree’ models:

\[ Y = g(z, x; T_1, M_1) + g(z, x; T_2, M_2) + \cdots + g(z, x; T_m, M_m) + \epsilon, \]
Bart and interactions

Data and fitted model; estimated treatment effect on treated units:
Summary of Chapter 20

- Splines and other weighted sums of basis functions
- Discrete or continuous parameterizations
- Multiple predictors
Summary of Chapter 20

- Splines and other weighted sums of basis functions
- Discrete or continuous parameterizations
- Multiple predictors
Summary of Chapter 20

- Splines and other weighted sums of basis functions
- Discrete or continuous parameterizations
- Multiple predictors
Summary of Chapter 20

- Splines and other weighted sums of basis functions
- Discrete or continuous parameterizations
- Multiple predictors
21. Gaussian process models

- Gaussian process regression
- Decomposition using a sum of Gaussian processes
- Logistic Gaussian processes
21. Gaussian process models

- Gaussian process regression
- Decomposition using a sum of Gaussian processes
- Logistic Gaussian processes
21. Gaussian process models

- Gaussian process regression
- Decomposition using a sum of Gaussian processes
- Logistic Gaussian processes
21. Gaussian process models

- Gaussian process regression
- Decomposition using a sum of Gaussian processes
- Logistic Gaussian processes
21.1. Gaussian process regression

- Mean function
- Covariance function
  - One possibility: \( c(x, x') = \phi_1 \exp(-\phi_2 ||x - x'||^2) \)
  - In \( d \) dimensions: \( c(x, x') = \phi_1 \exp \left( -\sum_{j=1}^{d} \alpha_j (x_j - x'_j)^2 \right) \)
- Gaussian process prior for coefficients of a basis expansion
- Computation using Gibbs, Metropolis, HMC
- Computation using expectation propagation
21.1. Gaussian process regression

- **Mean function**
  - Covariance function
    - One possibility: \( c(x, x') = \phi_1 \exp(-\phi_2 \|x - x'\|^2) \)
    - In \( d \) dimensions: \( c(x, x') = \phi_1 \exp \left( -\sum_{j=1}^{d} \alpha_j (x_j - x'_j)^2 \right) \)
  - Gaussian process prior for coefficients of a basis expansion
  - Computation using Gibbs, Metropolis, HMC
  - Computation using expectation propagation

Andrew Gelman
Bayesian Data Analysis, class 11b
21.1. Gaussian process regression

- **Mean function**
- **Covariance function**
  - One possibility: \( c(x, x') = \phi_1 \exp(-\phi_2 \|x - x'\|^2) \)
  - In \( d \) dimensions: \( c(x, x') = \phi_1 \exp \left(-\sum_{j=1}^{d} \alpha_j (x_j - x'_j)^2 \right) \)
- Gaussian process prior for coefficients of a basis expansion
- Computation using Gibbs, Metropolis, HMC
- Computation using expectation propagation
21.1. Gaussian process regression

- Mean function
- Covariance function
  - One possibility: $c(x, x') = \phi_1 \exp(-\phi_2 ||x - x'||^2)$
  - In $d$ dimensions: $c(x, x') = \phi_1 \exp \left( - \sum_{j=1}^{d} \alpha_j (x_j - x'_j)^2 \right)$
- Gaussian process prior for coefficients of a basis expansion
- Computation using Gibbs, Metropolis, HMC
- Computation using expectation propagation
Mean function

Covariance function

One possibility: 

\[ c(x, x') = \phi_1 \exp(-\phi_2 ||x - x'||^2) \]

In \( d \) dimensions:

\[ c(x, x') = \phi_1 \exp\left(-\sum_{j=1}^{d} \alpha_j (x_j - x'_j)^2\right) \]

Gaussian process prior for coefficients of a basis expansion

Computation using Gibbs, Metropolis, HMC

Computation using expectation propagation
21.1. Gaussian process regression

- Mean function
- Covariance function
  - One possibility: \( c(x, x') = \phi_1 \exp(-\phi_2 ||x - x'||^2) \)
  - In \( d \) dimensions: \( c(x, x') = \phi_1 \exp\left(-\sum_{j=1}^{d} \alpha_j (x_j - x'_j)^2\right) \)
- Gaussian process prior for coefficients of a basis expansion
  - Computation using Gibbs, Metropolis, HMC
  - Computation using expectation propagation
21.1. Gaussian process regression

- Mean function
- Covariance function
  - One possibility: \( c(x, x') = \phi_1 \exp(-\phi_2 ||x - x'||^2) \)
  - In \( d \) dimensions: \( c(x, x') = \phi_1 \exp \left( - \sum_{j=1}^{d} \alpha_j (x_j - x_j')^2 \right) \)
- Gaussian process prior for coefficients of a basis expansion
- Computation using Gibbs, Metropolis, HMC
- Computation using expectation propagation
21.1. Gaussian process regression

- Mean function
- Covariance function
  - One possibility: \( c(x, x') = \phi_1 \exp(-\phi_2 \|x - x'\|^2) \)
  - In \( d \) dimensions: \( c(x, x') = \phi_1 \exp \left( -\sum_{j=1}^{d} \alpha_j (x_j - x'_j)^2 \right) \)
- Gaussian process prior for coefficients of a basis expansion
- Computation using Gibbs, Metropolis, HMC
- Computation using expectation propagation
Draws from Gaussian process priors

Figure 21.1 Random draws from the Gaussian process prior with squared exponential covariance function showing different values of the amplitude parameter $\tau$ and the characteristic length scale parameter $l$. 
Figure 21.2 Posterior draws of a Gaussian process $\mu(x)$ fit to ten data points, conditional on three different choices of the parameters $\tau, l$ that characterize the process. Compare to Figure 21.1, which
Bayes inference with unknown amplitude and scale parameters

Figure 21.3 Marginal posterior distributions for Gaussian process parameters $\tau$, $l$ and error scale $\sigma$, and posterior mean and 90% region for $\mu(x)$, given the same ten data points from Figure 21.2.
21.2. Example: birthdays and birthdates

- A striking pattern
- Looking at more data
- Decomposition using a sum of Gaussian processes
- Model checking and improvement
21.2. Example: birthdays and birthdates

- A striking pattern
- Looking at more data
- Decomposition using a sum of Gaussian processes
- Model checking and improvement
21.2. Example: birthdays and birthdates

- A striking pattern
- Looking at more data
  - Decomposition using a sum of Gaussian processes
  - Model checking and improvement
21.2. Example: birthdays and birthdates

- A striking pattern
- Looking at more data
- Decomposition using a sum of Gaussian processes
- Model checking and improvement
21.2. Example: birthdays and birthdates

- A striking pattern
- Looking at more data
- Decomposition using a sum of Gaussian processes
- Model checking and improvement
A striking pattern in birthdates

Valentine’s Day: Two-Week Window

Halloween: Two-Week Window

* $p < .001$
Looking at all the days at once
A decomposition using Gaussian processes

Relative Number of Births in USA

Trends
- - - Slow trend
- Fast non-periodic component

Day of week effect
- * - 1972
- - - 1980
- - - - 1988

Seasonal effect
- * - 1972
- - - 1980
- - - - 1988

Day of year effect
- * - Valentine’s day
- - - Leap day
- - - April 1st
- - - Memorial day
- - - Independence day
- - - Labor day
- - - Thanksgiving
- - - Christmas
- - - New year
The model

\[ y_t(t) = f_1(t) + f_2(t) + f_3(t) + f_4(t) + f_5(t) + \epsilon_t \]

- **Long-term trends:**
  \[ f_1(t) \sim GP(0, c_1), \quad c_1(t, t') = \sigma_1^2 \exp \left(-\frac{|t-t'|^2}{l_1^2}\right) \]

- **Short-term variation:**
  \[ f_2(t) \sim GP(0, c_2), \quad c_2(t, t') = \sigma_2^2 \exp \left(-\frac{|t-t'|^2}{l_2^2}\right) \]

- **Weakly quasi-periodic:**
  \[ f_3(t) \sim GP(0, c_3), \quad c_3(t, t') = \sigma_3^2 \exp \left(-2 \sin^2 \left(\frac{\pi (t-t')}{7}\right)\right) \exp \left(-\frac{|t-t'|^2}{l_3^2}\right) \]

- **Yearly smooth seasonal:**
  \[ f_4(t) \sim GP(0, c_4), \quad c_4(s, s') = \sigma_4^2 \exp \left(-2 \sin^2 \left(\frac{\pi (s-s')}{365.25}\right)\right) \exp \left(-\frac{|s-s'|^2}{l_4^2}\right) \]

- **13 pre-chosen special days:**
  \[ f_5(t) = I_{\text{special day}}(t) \beta_a + I_{\text{weekend}}(t)I_{\text{special day}}(t) \beta_b \]

- **Unstructured residual:**
  \[ \epsilon_t \sim N(0, \sigma^2) \]

Now look for problems ...
The model

\[ y_t(t) = f_1(t) + f_2(t) + f_3(t) + f_4(t) + f_5(t) + \epsilon_t \]

- Long-term trends:
  \[ f_1(t) \sim \text{GP}(0, c_1), \quad c_1(t, t') = \sigma_1^2 \exp\left(-\frac{|t-t'|^2}{l_1^2}\right) \]

- Short-term variation:
  \[ f_2(t) \sim \text{GP}(0, c_2), \quad c_2(t, t') = \sigma_2^2 \exp\left(-\frac{|t-t'|^2}{l_2^2}\right) \]

- Weakly quasi-periodic:
  \[ f_3(t) \sim \text{GP}(0, c_3), \quad c_3(t, t') = \sigma_3^2 \exp\left(-\frac{2\sin^2(\pi(t-t')/7)}{l_{3,1}^2}\right) \exp\left(-\frac{|t-t'|^2}{l_{3,2}^2}\right) \]

- Yearly smooth seasonal:
  \[ f_4(t) \sim \text{GP}(0, c_4), \quad c_4(s, s') = \sigma_4^2 \exp\left(-\frac{2\sin^2(\pi(s-s')/365.25)}{l_{4,1}^2}\right) \exp\left(-\frac{|s-s'|^2}{l_{4,2}^2}\right) \]

- 13 pre-chosen special days:
  \[ f_5(t) = I_{\text{special day}}(t) \beta_a + I_{\text{weekend}}(t) I_{\text{special day}}(t) \beta_b \]

- Unstructured residual:
  \[ \epsilon_t \sim \text{N}(0, \sigma^2) \]

Now look for problems...
The model

\[ y_t(t) = f_1(t) + f_2(t) + f_3(t) + f_4(t) + f_5(t) + \epsilon_t \]

- Long-term trends:
  \[ f_1(t) \sim \text{GP}(0, c_1), \quad c_1(t, t') = \sigma_1^2 \exp \left( -\frac{|t-t'|^2}{l_1^2} \right) \]

- Short-term variation:
  \[ f_2(t) \sim \text{GP}(0, c_2), \quad c_2(t, t') = \sigma_2^2 \exp \left( -\frac{|t-t'|^2}{l_2^2} \right) \]

- Weakly quasi-periodic:
  \[ f_3(t) \sim \text{GP}(0, c_3), \quad c_3(t, t') = \sigma_3^2 \exp \left( -\frac{2 \sin^2(\pi(t-t')/7)}{l_{3,1}^2} \right) \exp \left( -\frac{|t-t'|^2}{l_{3,2}^2} \right) \]

- Yearly smooth seasonal:
  \[ f_4(t) \sim \text{GP}(0, c_4), \quad c_4(s, s') = \sigma_4^2 \exp \left( -\frac{2 \sin^2(\pi(s-s')/365.25)}{l_{4,1}^2} \right) \exp \left( -\frac{|s-s'|^2}{l_{4,2}^2} \right) \]

- 13 pre-chosen special days:
  \[ f_5(t) = I_{\text{special day}}(t) \beta_a + I_{\text{weekend}}(t) I_{\text{special day}}(t) \beta_b \]

- Unstructured residual:
  \[ \epsilon_t \sim \text{N}(0, \sigma^2) \]

Now look for problems...
The model

\[ y_t(t) = f_1(t) + f_2(t) + f_3(t) + f_4(t) + f_5(t) + \epsilon_t \]

- **Long-term trends:**
  \[ f_1(t) \sim \text{GP}(0, c_1), \quad c_1(t, t') = \sigma_1^2 \exp\left(-\frac{|t-t'|^2}{l_1^2}\right) \]

- **Short-term variation:**
  \[ f_2(t) \sim \text{GP}(0, c_2), \quad c_2(t, t') = \sigma_2^2 \exp\left(-\frac{|t-t'|^2}{l_2^2}\right) \]

- **Weakly quasi-periodic:**
  \[ f_3(t) \sim \text{GP}(0, c_3), \quad c_3(t, t') = \sigma_3^2 \exp\left(-\frac{2 \sin^2(\pi(t-t')/7)}{l_{3,1}^2}\right) \exp\left(-\frac{|t-t'|^2}{l_{3,2}^2}\right) \]

- **Yearly smooth seasonal:**
  \[ f_4(t) \sim \text{GP}(0, c_4), \quad c_4(s, s') = \sigma_4^2 \exp\left(-\frac{2 \sin^2(\pi(s-s')/365.25)}{l_{4,1}^2}\right) \exp\left(-\frac{|s-s'|^2}{l_{4,2}^2}\right) \]

- **13 pre-chosen special days:**
  \[ f_5(t) = I_{\text{special day}}(t) \beta_a + I_{\text{weekend}}(t) I_{\text{special day}}(t) \beta_b \]

- **Unstructured residual:**
  \[ \epsilon_t \sim \text{N}(0, \sigma^2) \]

- Now look for problems...
The model

\( y_t(t) = f_1(t) + f_2(t) + f_3(t) + f_4(t) + f_5(t) + \epsilon_t \)

- **Long-term trends:**
  \( f_1(t) \sim \text{GP}(0, c_1), \quad c_1(t, t') = \sigma_1^2 \exp \left( -\frac{|t-t'|^2}{l_1^2} \right) \)

- **Short-term variation:**
  \( f_2(t) \sim \text{GP}(0, c_2), \quad c_2(t, t') = \sigma_2^2 \exp \left( -\frac{|t-t'|^2}{l_2^2} \right) \)

- **Weakly quasi-periodic:**
  \( f_3(t) \sim \text{GP}(0, c_3), \quad c_3(t, t') = \sigma_3^2 \exp \left( -\frac{2 \sin^2(\pi (t-t')/7)}{l_{3,1}^2} \right) \exp \left( -\frac{|t-t'|^2}{l_{3,2}^2} \right) \)

- **Yearly smooth seasonal:**
  \( f_4(t) \sim \text{GP}(0, c_4), \quad c_4(s, s') = \sigma_4^2 \exp \left( -\frac{2 \sin^2(\pi (s-s')/365.25)}{l_{4,1}^2} \right) \exp \left( -\frac{|s-s'|^2}{l_{4,2}^2} \right) \)

- **13 pre-chosen special days:**
  \( f_5(t) = I_{\text{special day}}(t) \beta_a + I_{\text{weekend}}(t) I_{\text{special day}}(t) \beta_b \)

- **Unstructured residual:**
  \( \epsilon_t \sim N(0, \sigma^2) \)

- **Now look for problems . . .**
The model

\[ y_t(t) = f_1(t) + f_2(t) + f_3(t) + f_4(t) + f_5(t) + \epsilon_t \]

- **Long-term trends:**
  \[ f_1(t) \sim \text{GP}(0, c_1), \quad c_1(t, t') = \sigma_1^2 \exp \left(-\frac{|t-t'|^2}{l_1^2} \right) \]

- **Short-term variation:**
  \[ f_2(t) \sim \text{GP}(0, c_2), \quad c_2(t, t') = \sigma_2^2 \exp \left(-\frac{|t-t'|^2}{l_2^2} \right) \]

- **Weakly quasi-periodic:**
  \[ f_3(t) \sim \text{GP}(0, c_3), \quad c_3(t, t') = \sigma_3^2 \exp \left(-\frac{2 \sin^2(\pi (t-t')/7)}{l_{3,1}^2} \right) \exp \left(-\frac{|t-t'|^2}{l_{3,2}^2} \right) \]

- **Yearly smooth seasonal:**
  \[ f_4(t) \sim \text{GP}(0, c_4), \quad c_4(s, s') = \sigma_4^2 \exp \left(-\frac{2 \sin^2(\pi (s-s')/365.25)}{l_{4,1}^2} \right) \exp \left(-\frac{|s-s'|^2}{l_{4,2}^2} \right) \]

- **13 pre-chosen special days:**
  \[ f_5(t) = I_{\text{special day}}(t) \beta_a + I_{\text{weekend}}(t) I_{\text{special day}}(t) \beta_b \]

- **Unstructured residual:**
  \[ \epsilon_t \sim N(0, \sigma^2) \]

- **Now look for problems ...**
The model

\[ y_t(t) = f_1(t) + f_2(t) + f_3(t) + f_4(t) + f_5(t) + \epsilon_t \]

- Long-term trends: 
  \[ f_1(t) \sim \text{GP}(0, c_1), \quad c_1(t, t') = \sigma_1^2 \exp \left( -\frac{|t-t'|^2}{l_1^2} \right) \]

- Short-term variation: 
  \[ f_2(t) \sim \text{GP}(0, c_2), \quad c_2(t, t') = \sigma_2^2 \exp \left( -\frac{|t-t'|^2}{l_2^2} \right) \]

- Weakly quasi-periodic: 
  \[ f_3(t) \sim \text{GP}(0, c_3), \quad c_3(t, t') = \sigma_3^2 \exp \left( -\frac{2 \sin^2(\pi(t-t')/7)}{l_{3,1}^2} \right) \exp \left( -\frac{|t-t'|^2}{l_{3,2}^2} \right) \]

- Yearly smooth seasonal: 
  \[ f_4(t) \sim \text{GP}(0, c_4), \quad c_4(s, s') = \sigma_4^2 \exp \left( -\frac{2 \sin^2(\pi(s-s')/365.25)}{l_{4,1}^2} \right) \exp \left( -\frac{|s-s'|^2}{l_{4,2}^2} \right) \]

- 13 pre-chosen special days: 
  \[ f_5(t) = \mathbb{I}_{\text{special day}}(t) \beta_a + \mathbb{I}_{\text{weekend}}(t) \mathbb{I}_{\text{special day}}(t) \beta_b \]

- Unstructured residual: 
  \[ \epsilon_t \sim \mathcal{N}(0, \sigma^2) \]

Now look for problems ....
The model

\( y_t(t) = f_1(t) + f_2(t) + f_3(t) + f_4(t) + f_5(t) + \epsilon_t \)

- **Long-term trends:**
  \( f_1(t) \sim \text{GP}(0, c_1), \quad c_1(t, t') = \sigma_1^2 \exp\left(-\frac{|t-t'|^2}{l_1^2}\right) \)

- **Short-term variation:**
  \( f_2(t) \sim \text{GP}(0, c_2), \quad c_2(t, t') = \sigma_2^2 \exp\left(-\frac{|t-t'|^2}{l_2^2}\right) \)

- **Weakly quasi-periodic:**
  \( f_3(t) \sim \text{GP}(0, c_3), \quad c_3(t, t') = \sigma_3^2 \exp\left(-\frac{2\sin^2(\pi(t-t')/7)}{l_{3,1}^2}\right) \exp\left(-\frac{|t-t'|^2}{l_{3,2}^2}\right) \)

- **Yearly smooth seasonal:**
  \( f_4(t) \sim \text{GP}(0, c_4), \quad c_4(s, s') = \sigma_4^2 \exp\left(-\frac{2\sin^2(\pi(s-s')/365.25)}{l_{4,1}^2}\right) \exp\left(-\frac{|s-s'|^2}{l_{4,2}^2}\right) \)

- **13 pre-chosen special days:**
  \( f_5(t) = I_{\text{special day}}(t)\beta_a + I_{\text{weekend}}(t)I_{\text{special day}}(t)\beta_b \)

- **Unstructured residual:**
  \( \epsilon_t \sim N(0, \sigma^2) \)

- **Now look for problems . . .**
Problems with the inferences?

Relative Number of Births in USA

- Slow trend
- Fast non-periodic component

Day of week effect

Seasonal effect

Day of year effect

Andrew Gelman  Bayesian Data Analysis, class 11b
Inferences from an improved model

Relative Number of Births in USA

Trends
- Slow trend
- Fast non-periodic component

Day of week effect
- 1972
- 1980
- 1988

Seasonal effect
- 1972
- 1980
- 1988

Day of year effect
- Valentine's day
- Leap day
- April 1st
- Memorial day
- Independence day
- Labor day
- Thanksgiving
- Christmas

Andrew Gelman
Bayesian Data Analysis, class 11b
21.3. Latent Gaussian process models

- Similar to a generalized linear model
- Normal approximation for computation: Laplace’s method and expectation propagation
- Survival data regression in a leukemia study:
21.3. Latent Gaussian process models

- Similar to a generalized linear model
- Normal approximation for computation: Laplace’s method and expectation propagation
- Survival data regression in a leukemia study:
  - Log-logistic model for survival times
  - Use cross-validation to demonstrate better fit with nonlinear model
21.3. Latent Gaussian process models

- Similar to a generalized linear model
- Normal approximation for computation: Laplace’s method and expectation propagation
- Survival data regression in a leukemia study:
  - Log-logistic model for survival times
  - Use cross-validation to demonstrate better fit with nonlinear model
21.3. Latent Gaussian process models

- Similar to a generalized linear model
- Normal approximation for computation: Laplace’s method and expectation propagation
- Survival data regression in a leukemia study:
  - Log-logistic model for survival times
  - Use cross-validation to demonstrate better fit with nonlinear model
21.3. Latent Gaussian process models

- Similar to a generalized linear model
- Normal approximation for computation: Laplace’s method and expectation propagation
- Survival data regression in a leukemia study:
  - Log-logistic model for survival times
  - Use cross-validation to demonstrate better fit with nonlinear model
21.3. Latent Gaussian process models

- Similar to a generalized linear model
- Normal approximation for computation: Laplace’s method and expectation propagation
- Survival data regression in a leukemia study:
  - Log-logistic model for survival times
  - Use cross-validation to demonstrate better fit with nonlinear model
Regression for survival data

Expected lifetime (days)

Age (years)

Expected lifetime (days)

WBC ($\log_{10}(50 \times 10^9 / L)$)

Expected lifetime (days)

Townsend deprivation index (TDI)

Expected lifetime (days)

WBC ($\log_{10}(50 \times 10^9 / L)$)
21.4. Functional data analysis

- Data are random functions
- Item $i$, observation time $j$ at time $t_{ij}$:

$$ y_{ij} \sim N(f_i(t_{ij}), \sigma^2) $$

- With predictors $x_i$:

$$ y_{ij} \sim N(f(x_i, t_{ij}), \sigma^2) $$

- Gaussian process prior $f \sim \text{GP}(m, c)$
- Cov matrix could have squared exponential form:

$$ \tau^2 \exp \left( - \sum_{j=1}^{p} \frac{(x_j - x_j')^2}{l_j^2} + \frac{(t - t')^2}{l_{p+1}^2} \right) $$
21.4. Functional data analysis

- Data are random functions
  - Item \( i \), observation time \( j \) at time \( t_{ij} \):
    \[
y_{ij} \sim N(f_i(t_{ij}), \sigma^2)
    \]
  - With predictors \( x_i \):
    \[
y_{ij} \sim N(f(x_i, t_{ij}), \sigma^2)
    \]
- Gaussian process prior \( f \sim \text{GP}(m, c) \)
- Cov matrix could have squared exponential form:
  \[
  \tau^2 \exp \left( - \sum_{j=1}^{p} \frac{(x_j - x'_j)^2}{l_j^2} + \frac{(t - t')^2}{l^2_{p+1}} \right)
  \]
21.4. Functional data analysis

- Data are random functions
- Item $i$, observation time $j$ at time $t_{ij}$:
  \[ y_{ij} \sim N(f_i(t_{ij}), \sigma^2) \]

- With predictors $x_i$:
  \[ y_{ij} \sim N(f(x_i, t_{ij}), \sigma^2) \]

- Gaussian process prior $f \sim GP(m, c)$
- Cov matrix could have squared exponential form:
  \[ \tau^2 \exp \left( - \sum_{j=1}^{p} \frac{(x_j - x_j')^2}{l_j^2} + \frac{(t - t')^2}{l_{p+1}^2} \right) \]
21.4. Functional data analysis

- Data are random functions
- Item $i$, observation time $j$ at time $t_{ij}$:
  \[ y_{ij} \sim N(f_i(t_{ij}), \sigma^2) \]
- With predictors $x_i$:
  \[ y_{ij} \sim N(f(x_i, t_{ij}), \sigma^2) \]
- Gaussian process prior $f \sim \text{GP}(m, c)$
- Cov matrix could have squared exponential form:
  \[
  \tau^2 \exp \left( - \sum_{j=1}^{p} \frac{(x_j - x'_j)^2}{l_j^2} + \frac{(t - t')^2}{l_{p+1}^2} \right)
  \]
21.4. Functional data analysis

- Data are random functions
- Item $i$, observation time $j$ at time $t_{ij}$:
  \[ y_{ij} \sim N(f_i(t_{ij}), \sigma^2) \]
- With predictors $x_i$:
  \[ y_{ij} \sim N(f(x_i, t_{ij}), \sigma^2) \]
- Gaussian process prior $f \sim \text{GP}(m, c)$
- Cov matrix could have squared exponential form:
  \[
  \tau^2 \exp \left( - \sum_{j=1}^{p} \frac{(x_j - x_j')^2}{l_j^2} + \frac{(t - t')^2}{l_{p+1}^2} \right)
  \]
21.4. Functional data analysis

- Data are random functions
- Item $i$, observation time $j$ at time $t_{ij}$:
  \[ y_{ij} \sim N(f_i(t_{ij}), \sigma^2) \]
- With predictors $x_i$:
  \[ y_{ij} \sim N(f(x_i, t_{ij}), \sigma^2) \]
- Gaussian process prior $f \sim GP(m, c)$
- Cov matrix could have squared exponential form:
  \[ \tau^2 \exp \left( - \sum_{j=1}^{p} \frac{(x_j - x_j')^2}{l_j^2} + \frac{(t - t')^2}{l_{p+1}^2} \right) \]
21.5. Density estimation and regression

- Logistic Gaussian process
- Simple examples of density estimation
- Density regression
- For small problems, compute in Stan
- For large problems, compute using Laplace’s method or expectation propagation
21.5. Density estimation and regression

- Logistic Gaussian process
- Simple examples of density estimation
- Density regression
- For small problems, compute in Stan
- For large problems, compute using Laplace’s method or expectation propagation
21.5. Density estimation and regression

- Logistic Gaussian process
- Simple examples of density estimation
- Density regression
- For small problems, compute in Stan
- For large problems, compute using Laplace’s method or expectation propagation
21.5. Density estimation and regression

- Logistic Gaussian process
- Simple examples of density estimation
- Density regression
  - For small problems, compute in Stan
  - For large problems, compute using Laplace’s method or expectation propagation
21.5. Density estimation and regression

- Logistic Gaussian process
- Simple examples of density estimation
- Density regression
- For small problems, compute in Stan
- For large problems, compute using Laplace’s method or expectation propagation
Logistic Gaussian process for density estimation

- Prior distribution for a probability density:
  - $f(y)$ is a Gaussian process
  - Density function:
    $$ p(y|f) = \frac{e^{f(y)}}{\int e^{f(y')} dy'} $$

- Alternative form:
  $$ p(y) = g_0(y) \frac{e^{W(G_0(y))}}{\int e^{W(v)} dv} $$

- $W(t)$ is a Gaussian process on $[0, 1]$
- $g_0(y)$ is a specified probability density function
Prior distribution for a probability density:

- $f(y)$ is a Gaussian process
- Density function:

$$p(y|f) = \frac{e^{f(y)}}{\int e^{f(y')} dy'}$$

Alternative form:

$$p(y) = g_0(y) \frac{e^{W(G_0(y))}}{\int e^{W(v)} dv}$$
Prior distribution for a probability density:

- \( f(y) \) is a Gaussian process
- Density function:

\[
p(y|f) = \frac{e^{f(y)}}{\int e^{f(y')} \, dy'}
\]

Alternative form:

\[
p(y) = g_0(y) \frac{e^{W(G_0(y))}}{\int e^{W(v)} \, dv}
\]

- \( W(t) \) is a Gaussian process on \([0, 1] \)
- \( g_0(y) \) is a specified probability density function
Logistic Gaussian process for density estimation

- Prior distribution for a probability density:
  - \( f(y) \) is a Gaussian process
  - Density function:
    \[
    p(y|f) = \frac{e^{f(y)}}{\int e^{f(y')} dy'}
    \]

- Alternative form:
  \[
  p(y) = g_0(y) \frac{e^{W(G_0(y))}}{\int e^{W(v)} dv}
  \]

- \( W(t) \) is a Gaussian process on \([0, 1]\)
- \( g_0(y) \) is a specified probability density function
Logistic Gaussian process for density estimation

- Prior distribution for a probability density:
  - \( f(y) \) is a Gaussian process
  - Density function:
    \[
    p(y|f) = \frac{e^{f(y)}}{\int e^{f(y')} dy'}
    \]

- Alternative form:
  \[
  p(y) = g_0(y) \frac{e^{W(G_0(y))}}{\int e^{W(v)} dv}
  \]
  - \( W(t) \) is a Gaussian process on \([0, 1]\)
  - \( g_0(y) \) is a specified probability density function
Logistic Gaussian process for density estimation

- Prior distribution for a probability density:
  - \( f(y) \) is a Gaussian process
  - Density function:
    \[
    p(y|f) = \frac{e^{f(y)}}{\int e^{f(y')} dy'}
    \]

- Alternative form:
  \[
  p(y) = g_0(y) \frac{e^{W(G_0(y))}}{\int e^{W(v)} dv}
  \]
  - \( W(t) \) is a Gaussian process on [0, 1]
  - \( g_0(y) \) is a specified probability density function
Logistic Gaussian process for density estimation

- Prior distribution for a probability density:
  - $f(y)$ is a Gaussian process
  - Density function:
    \[
    p(y|f) = \frac{e^{f(y)}}{\int e^{f(y')} dy'}
    \]

- Alternative form:
  \[
  p(y) = g_0(y) \frac{e^{W(G_0(y))}}{\int e^{W(v)} dv}
  \]

- $W(t)$ is a Gaussian process on $[0, 1]$  
- $g_0(y)$ is a specified probability density function
Acidity data and galaxy data
Summary of Chapter 21

- Gaussian processes: discrete or continuous parameterizations
- Nonlinear link functions
- Computing
Summary of Chapter 21

- Gaussian processes: discrete or continuous parameterizations
- Nonlinear link functions
- Computing
Summary of Chapter 21

- Gaussian processes: discrete or continuous parameterizations
- Nonlinear link functions
- Computing
Summary of Chapter 21

- Gaussian processes: discrete or continuous parameterizations
- Nonlinear link functions
- Computing
Homework due beginning of class 12b

- All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf
  - Computing problem: Bayesian inference for a spline model
  - Computing problem: Bayesian inference for a Gaussian process model
  - Applied problem: Social networks and gay marriage
Homework due beginning of class 12b

- All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf
  - Computing problem: Bayesian inference for a spline model
  - Computing problem: Bayesian inference for a Gaussian process model
  - Applied problem: Social networks and gay marriage
Homework due beginning of class 12b

- All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf
  - Computing problem: Bayesian inference for a spline model
  - Computing problem: Bayesian inference for a Gaussian process model
  - Applied problem: Social networks and gay marriage
Homework due beginning of class 12b

- All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf
  - Computing problem: Bayesian inference for a spline model
  - Computing problem: Bayesian inference for a Gaussian process model
  - Applied problem: Social networks and gay marriage
Homework due beginning of class 12b

- All assignments are at http://www.stat.columbia.edu/~gelman/bda.course/homeworks.pdf
  - Computing problem: Bayesian inference for a spline model
  - Computing problem: Bayesian inference for a Gaussian process model
  - Applied problem: Social networks and gay marriage