Gaussian Process for Spatio-temporal Processes, Inverse Problem by Sampling

Adway Mitra

1 February 2021

Multivariate Gaussian Distribution

- Consider random variables $X_1, X_2, \dots, X_D$
- $X = \{X_1, X_2, \dots, X_D\}$
- $p(X) = \frac{1}{(2\pi)^{D/2} det(\Sigma)} exp(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu))$
- Σ is the $D \times D$ covariance matrix, $\Sigma(i,j) = Cov(X_i, X_j)$
- Covariance matrix size blows up with D!
- Possible solution: replace covariance matrix by covariance function!

Gaussian Process

- Consider a (finite or infinite) set of random variables X_1, X_2, ...
- Suppose each of them represents a location
- Consider any random finite subset $\{X_{i1}, X_{i2}, \ldots, X_{iN}\}$
- Then we have $(X_{i1}, X_{i2}, \ldots, X_{iN}) \sim \mathcal{N}(\mu, \Sigma)$
- $\mu(s)$: mean function (a function of s)
- $\Sigma(s, s') = K(||s - s'||)$, where K is the covariance function (a function of $||s - s'||$)

Gaussian Process

- $X \sim \mathcal{GP}(\mu, K)$, where $\mathcal{GP}$ represents Gaussian Process, with K as the covariance function
- X is now a continuous function, defined at every location
- Any finite set of locations follows multivariate Gaussian distribution
- Relationship between X at different locations represented through covariance function
- Given observations at each location, we can predict the values at other locations

Interpolation using Gaussian Processes

- Consider any location $s+1$ where we want to predict X_{s+1}
- Conditional PDF $p(X_{s+1}|X_1,\ldots,X_s) = \frac{p(X_1,X_2,\ldots,X_s,X_{s+1})}{p(X_1,X_2,\ldots,X_s)}$
- Both joint PDFs $p(X_1, X_2, \ldots, X_s, X_{s+1})$ and $p(X_1, X_2, \ldots, X_s)$ are Gaussian!
- For both, mean vector and covariance matrix will come from the GP mean and covariance functions!

Spatio-temporal Hierarchical Gaussian Process

- Data Model: $X \sim \mathcal{N}(\mu + \eta, \sigma I)$
- $\eta = AZ + BY$
- $\mu \sim \mathcal{GP}(\mu_0, K_0)$?
- $Z \sim \mathcal{GP}(\mu_Z, K_Z)$?
- These Gaussian Processes can be either spatial or temporal
- We can decompose μ and Z into spatial and temporal components

Spatio-temporal Hierarchical Gaussian Process

- Data Model: $X(s,t) \sim \mathcal{N}(\mu(s,t) + \eta(s,t), \sigma)$
- Parameter Model: $\mu(s,t) = \mu_S(s)\mu_T(t)$
- $\mu_S \sim \mathcal{GP}(\mu_{S0}, K_{S0})$, $\mu_T \sim \mathcal{GP}(\mu_{T0}, K_{T0})$
- Process Model: $\eta(s,t) = AZ(s,t) + BY(s,t)$ where $Z(s,t) = Z_S(s)Z_T(t)$
- $Z_S \sim \mathcal{GP}(\mu_{S0}, K_{S0})$, $\mu_T \sim \mathcal{GP}(\mu_{T0}, K_{T0})$

Forward Problem: Data Generation

- Forward problem: generate/simulate X using this model
- Identify a set of locations and time-points to generate data
- Generate $\mu(s,t)$ from the Gaussian Processes μ_S and μ_T
- For each of μ_S and μ_T, first generate data at one location/time-point
- Keep sampling as $p(\mu_s|\mu_1,\ldots,\mu_{s-1})$ and $p(\mu_t|\mu_1,\ldots,\mu_{t-1})$
- Similarly generate $Z(s,t)$ from the Gaussian Processes Z_S and Z_T
- Finally generate $X(s,t)$ using the data model

Inverse Problem

- We already have the observation X and covariates Y
- Can we estimate Z, μ, σ, A, B etc?
- Z, μ are latent random variables, σ, A, B are parameters
- Let us assume that parameters are fixed and known
- We need to estimate Z and μ by *Gibbs Sampling*
- Assign initial values to all Z and μ variables
- Optimal values to be found by repeated sampling of each variable

Inverse Problem

- Sample a new value of $Z(s,t)$ from $p(Z(s,t)|Z,X,\mu)$, conditional distribution of $Z(s,t)$ based on all other variables
- Repeat this process for all $s \in \{1,S\}$ and $t \in \{1,T\}$
- Sample a new value of $\mu(s,t)$ from $p(\mu(s,t)|\mu,X,Z)$, conditional distribution of $\mu(s,t)$ based on all other variables
- Repeat this process for all $s \in \{1,S\}$ and $t \in \{1,T\}$
- Store the sampled values of each variable
- Repeat the whole process for many iterations
- For each variable, select the mode of its samples

ASSIGNMENT (10%)

Generate spatio-temporal data over a 20×20 grid system for 100 time-steps

1. Generate μ using a Gaussian Process (μ_S, μ_T) using suitable covariance functions
2. Generate Z using a different Gaussian Process
3. Select the covariates Y in your own way
4. Choose A, B, σ and generate X
5. Calculate the variogram between different pairs of locations and plot histogram as a function of distance