Gaussian Process Latent Variable Model (for WiFi-SLAM) in 1-D

Project Presentation

Imran Shafique Ansari

02\textsuperscript{nd} August 2011
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
</tr>
<tr>
<td>2</td>
<td>Gaussian Process Latent Variable Models</td>
</tr>
<tr>
<td>3</td>
<td>GP-LVM for WiFi SLAM</td>
</tr>
<tr>
<td>4</td>
<td>Conclusion</td>
</tr>
</tbody>
</table>
Introduction

- WiFi Localization → determining location from wireless signal strengths
- A novel technique solving WiFi-SLAM using GP-LVM
- To determine latent-space locations of unlabeled signal strength data
- ∴ employ GP-LVM with motion model to localize efficiently with help of signal strength data
- Basically, GP-LVM maps high-dimensional data to low-dimensional data
Introduction

Constraints:

- Locations → Locations
- Distance between successive positions

Related Work ¹

Let $D = (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be our data and

$$y_i = f(x_i) + \epsilon,$$

be our noisy process where each $x_i$ is an input sample in $\mathbb{R}^q$, each $y_i$ is a target value or observation in $\mathbb{R}$, and $\epsilon$ is assumed to be zero mean, additive Gaussian distribution with variance $\sigma_n^2$. A key requirement of GPs is correlation between the different points. This is assumed as

$$\text{cov}(f(x_i), f(x_j)) = k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{1}{2l^2}|x_i - x_j|^2\right),$$

where $\sigma_f^2$ is the signal variance and $l$ is a length scale that determines the strength of correlation between points.
This can be additionally presented as follows in presence of noise

\[ \text{cov}(y_r, y_j) = k(x_i, x_j) + \sigma_n^2 \delta_{rs}, \]

where \( \delta_{rs} \) is one for \( r = s \) else zero. The above equation may also be given as

\[ \text{cov}(Y) = K + \sigma_n^2 I. \]
It follows that the posterior over function values is Gaussian:

\[ p(f(x^*_\star)|x^*_\star, X, Y) = \mathcal{N}(f(x^*_\star); \mu_{x^*_\star}, \sigma_{x^*_\star}^2), \]

where \( \mu_{x^*_\star} = k^T_\star (K + \sigma_n^2 I)^{-1} Y \) and \( \sigma_{x^*_\star}^2 = k(x^*_\star, x^*_\star) - k^T_\star (K + \sigma_n^2 I)^{-1} k_\star. \)

Now to achieve the final result, one needs to include the noise as

\[ p(y^*_\star|x^*_\star, X, Y) = \int p(y^*_\star|f(x^*_\star))p(f(x^*_\star)|x^*_\star, X, Y)df(x^*_\star) = \mathcal{N}(y^*_\star; \mu_{x^*_\star}, \sigma_{x^*_\star}^2 + \sigma_n^2). \]
Known Locations

Results

Figure: Actual vs. Random Positions.
Model:

\[ y_{ij} = f(x_i; w_j) + \epsilon \]

Assuming independence,

\[ p(Y|X, W) = \prod_{ij} p(y_{ij}|x_i, w_j). \]

Further assuming Gaussian components and maximizing the above likelihood with respect to X, following is obtained

\[ p(Y|X) = \prod_j \mathcal{N}(y_j; 0, XX^T + \sigma_n^2 I). \]

This is the expression that can be utilized with a linear covariance matrix, in place of the variance currently obtained above, suitng our particular needs.
The likelihood of the model to be considered in this work takes the form as

\[ p(X, Y) = p(Y|X)p(X), \]

where \( p(Y|X) \) is modeled by a GP and \( p(X) \) is the expression of our latent-space dynamics model. \( p(X) \) is given based on the constraint mentioned earlier in introduction section as

\[ p(X) = p(\text{distance}). \]
This constrain probability is given by

\[ p(\text{distance}) = \prod_{i=1}^{n-1} \mathcal{N}(\|x_{i+1} - x_i\|; \Delta_i \mu_v, \Delta_i \sigma_v) , \]

where \( x_i \)'s are the rows of \( X \), \( \Delta_i \) is time step, \( \mu_v \) is velocity mean, and \( \sigma_v \) is velocity variance.

Also, \( X \) is \( \mathbb{R}^{n \times q} \) and \( Y \) is \( \mathbb{R}^{n \times d} \) with \( n \) as number of robots, \( q \) is the number of steps taken, and \( d \) is the number of Wi-Fi routers.
Analytical Results

\[
p(Y|X) = \frac{1}{[2\pi^{d/2}|XX^T + \sigma_n^2 I|^{1/2}]^n} \exp(-\frac{1}{2}(XX^T + \sigma_n^2 I)^{-1}(y_1^T y_1 + y_2^T y_2 + \cdots + y_n^T y_n))
\]

\[
p(X) = \frac{1}{(2\pi\Delta_1 \sigma_v^2)^{1/2} \cdots (2\pi\Delta_{n-1} \sigma_v^2)^{1/2}} \exp\left(-\frac{1}{2\sigma_v^2} \left( \frac{[(X^T B_{2,1} X)^{1/2} - \Delta_1 \mu_v]^2}{\Delta_1} \right) + \cdots + \frac{[(X^T B_{n,n-1} X)^{1/2} - \Delta_{n-1} \mu_v]^2}{\Delta_{n-1}} \right)
\]
Applying Negative log-likelihood,

$$\log p(Y|X) = C - \frac{n}{2} \log |XX^T + \sigma_n^2 I| - \frac{1}{2} \sum_{j=1}^{d} y_j^T (XX^T + \sigma_n^2 I)^{-1} y_j$$

$$- \frac{1}{2\sigma_v^2} \sum_{i=1}^{n-1} \frac{1}{\Delta_i} \left[ (X^T B_{i+1,i}^T X)^{1/2} - \Delta_i \mu_v \right]^2,$$

where $C$ is a constant term and $B_{i+1,i} = A_{i+1} - A_i$ with $A = I$ of size $n$. 
Now, applying gradient-descent,

From the reference given below, we have,

$$\nabla \log |XX^T + \sigma_n^2 I| = (XX^T + \sigma_n^2 I)^{-1} X.$$

Therefore,

$${\nabla \log p(Y|X) = C - n(XX^T + \sigma_n^2 I)^{-1} X + \left( \sum_{j=1}^{d} y_j^T (XX^T + \sigma_n^2 I)^{-2} y_j \right) X - \frac{1}{\sigma_v^2} \sum_{i=1}^{n-1} \frac{1}{\Delta_i} [(X^T B_{i+1, i} X)^{1/2} - \Delta_i \mu_v] (X^T B_{i+1, i} X)^{-1/2} (B_{i+1, i} X).}$$

---

Simulation Results

Figure: Log-likelihood Plot.
Figure: Log-likelihood Contour Plot.
Figure: Log-likelihood plot.
Finally, with these definitions, the marginalized latent positions $X$ will be optimized by minimizing the negative log-likelihood of the full GP-LVM likelihood model $P(X, Y)$ obtained using gradient-descent algorithm.
THANK YOU