Teleinformatics Engineering Department, Federal University of Ceará

Introduction to Gaussian Processes

Filipe P. de Farias, IC
filipepfarias@fisica.ufc.br

October 11, 2018

Preamble

The **Gaussian Processes** are the widely used stochastic processes for modeling dependent data observed over time, space or even time and space. Here, we'll iniciate our study with a **Probability and Random Process Theory Review** taking some point to base our journey, going through **Linear Regression** and finally the **Gaussian Processes**.

The material here presented isn't sufficient to guide you over basic probability, so it's recommended to have some knowledge, once we'll just take a simple review.

Outline

1 Probability and Random Process Theory Review
1.1 Basic Concepts of Probability Theory

2 Linear Regression
2.1 Basic Concepts of Curve Fitting
2.2 Bayesian Curve Fitting

Probability and Random Process Theory Review

Specifying Random Experiments

Let's start with the sentence

Sentence

*A **random experiment** is specified by stating an **experimental procedure** and a **set** of one or more measurements or observations.*

Overleaf users

Warning

You can ignore this slide if you're **not** working with Overleaf.

Overleaf, Beamer and Biber do not always get along well together. For this reason, if you make a mistake while writing this presentation, in the drop-down error message you'll **always** get Biber-related error messages.

Luckily, you just have to click on "`go to first error/warning`" and the UI will scroll to the line containing your mistake.

Compiling

Warning

You can ignore this slide if you're working with Overleaf.

To compile this deck you'll need the `biber` package. Probably your TEXeditor already supports it; if not, you will easily find online the instructions to install it.

If you're not using an editor, you can compile this presentation using the command line by running:

```
$ pdflatex main.tex
$ biber main.bcf
$ pdflatex main.tex
$ pdflatex main.tex
```

Linear Regression

Bayesian Curve Fitting

So, we'll start to look the regression with a statistical approach. To encourage you, let's take the sentence.

So, we'll start to look the regression with a statistical approach. To encourage you, let's take the sentence.

Sentence

If we could update the ***regression weights*** *as we acquire some new values of the experiment?*

Bayesian Curve Fitting

Let's take a look again at the Bayes Theorem

Let's take a look again at the Bayes Theorem

Bayes Theorem

$$p(\mathbf{w}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathbf{w})\,\overbrace{p(\mathbf{w})}^{\text{the } \textit{a priori} \text{ probability}}}{\underbrace{p(\mathcal{D})}_{\text{the } \textit{a priori} \text{ probability}}} \tag{1}$$

Bayesian Curve Fitting

Let's take a look again at the Bayes Theorem

Bayes Theorem

$$p(\mathbf{w}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathbf{w})\,\overbrace{p(\mathbf{w})}^{\text{the } a \text{ } priori \text{ probability}}}{\underbrace{p(\mathcal{D})}_{\text{the } a \text{ } priori \text{ probability}}} \tag{1}$$

So, if **we have the probability** of the data, we'll could estimate the **future weights**.

Let's take a look again at the Bayes Theorem

Bayes Theorem

$$p(\mathbf{w}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathbf{w})\overbrace{p(\mathbf{w})}^{\text{the } a \text{ } priori \text{ probability}}}{\underbrace{p(\mathcal{D})}_{\text{the } a \text{ } priori \text{ probability}}} \tag{1}$$

So, if **we have the probability** of the data, we'll could estimate the **future weights**.

But, how?

Bayesian Curve Fitting

Taking some steps back, let's re-visit the **Curve Fitting**. There, the strategy was minimize the error function.

Bayesian Curve Fitting

Taking some steps back, let's re-visit the **Curve Fitting**. There, the strategy was minimize the error function.

Now we'll try to view the same problem with a *probabilistic perspective*. We're trying to make predictions for the target value $\boldsymbol{t}$ given some new values of x.

Bayesian Curve Fitting

Taking some steps back, let's re-visit the **Curve Fitting**. There, the strategy was minimize the error function.

Now we'll try to view the same problem with a *probabilistic perspective*. We're trying to make predictions for the target value $\boldsymbol{t}$ given some new values of x.

A good ideia is to express our target values $\boldsymbol{t}$ in terms of **gaussians distributions** with the mean equals to $y(x, \mathbf{w})$.

Bayesian Curve Fitting

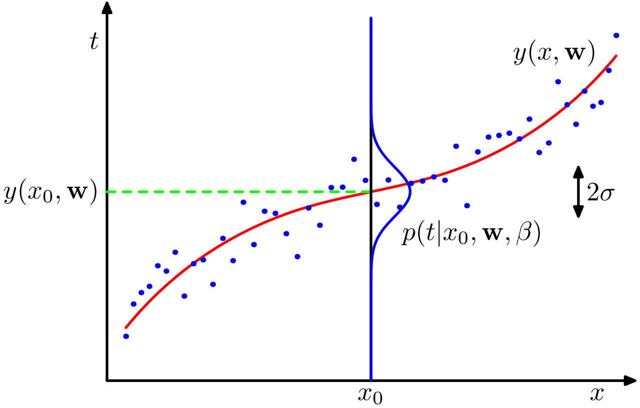

Figure: Schematic of the polynomial function $y(x, \mathbf{w})$ and the gaussian distribution p.

Title

$$blah = blah = \begin{cases} blah \\ blah \end{cases}$$

Title

$blah = result$