Statistical Learning Survey

Abraham Neuwirth

May 29, 2017

Contents

1 Introduction

Christopher Bishop (Bishop 2006) in his seminal book "Pattern Recognition and Machine Learning", introduces the field of statistical learning with a classical story:

> [T]he extensive astronomical observations of Tycho Brahe in the 16th century allowed Johannes Kepler to discover the empirical laws of planetary motion, which in turn provided a springboard for the development of classical mechanics.

It's the archtype of statistical learning success stories. Lot's of data, brilliant minds, and a model to illuminate and explain it all.

2 Supervised Learning

Lemma 2.1. *This is a lemma*

2.1 Least-Squares Regression

2.2 Support Vector Machines

3 Unsupervised Learning

3.1 Classification

3.2 Naive Bayes

3.3 K-Nearest Neighbors

4 Regularlization

Overfitting is a problem. regularlization penalizes biggest predictors. "Regularization can be accomplished by restricting the hypothesis space $\mathcal{H}$"

4.1 Tikhonov regularization (Ridge Regression)

4.2 Lasso Regression

4.3 Principal Components

Dimensionality reduciton

5 Reinforcment Learning

6 Deep Learning

This is a level 1 heading

6.1 Level 2

Blah blah (se also James et al. 2014 ch. 1) and Hastie, Tibshirani, and Friedman (2001).

6.1.1 equations

This s an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com. See equation (6.1) there's also (6.2)

$$\sum_{i}^{x} 5 + 1 \tag{6.1}$$

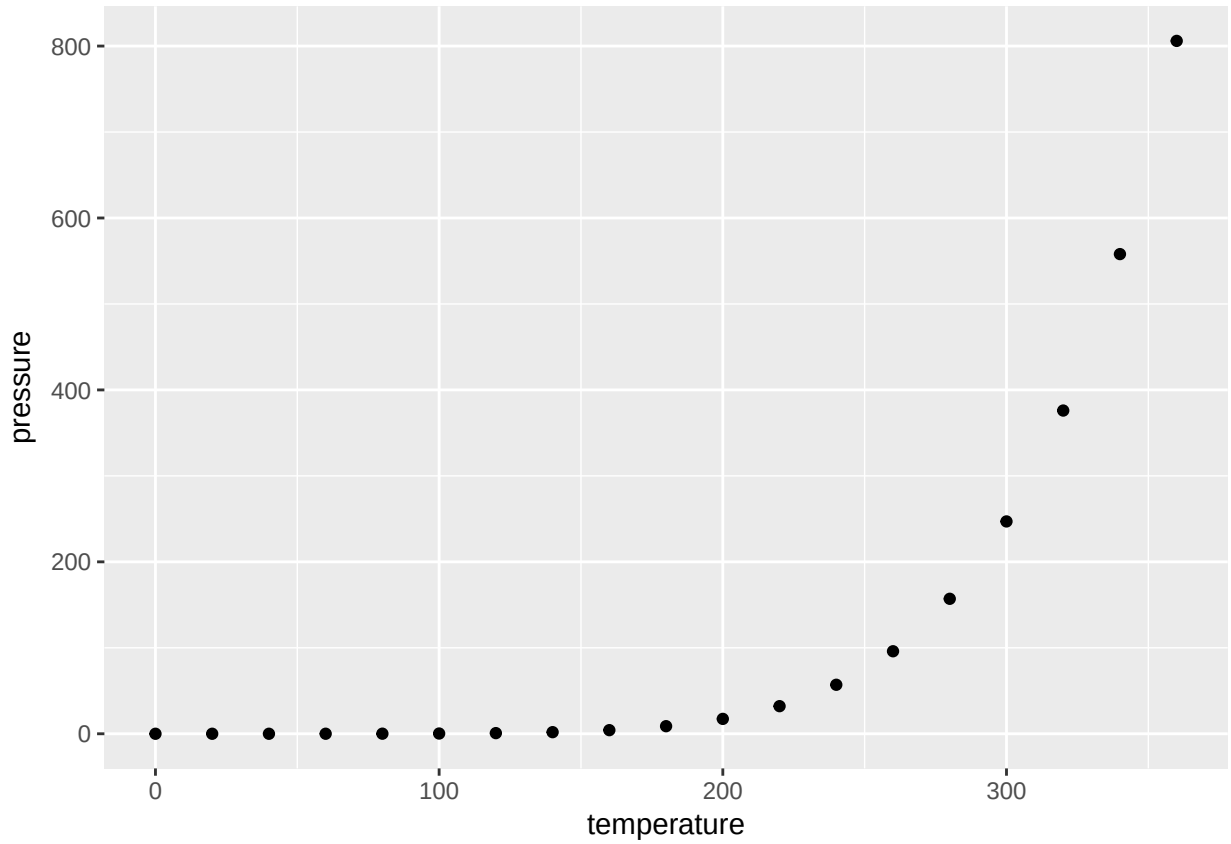

Figure 6.1: test a caption

$$\sum_{i}^{x} 5 + 5 \tag{6.2}$$

6.1.2 Level 4

Theorem 6.1. *Here is A theorem.*

6.2 Including Plots

You can also embed plots, for example:

Note that the `echo = FALSE` parameter was added to the code chunk 6.1 to prevent printing of the R code that generated the plot. Bishop (2006)

also, lets put this in here

$$
\begin{aligned}
\text{Var}(\hat{\beta}) &= \text{Var}((X'X)^{-1}X'y) \\
&= (X'X)^{-1}X'\text{Var}(y)((X'X)^{-1}X')' \\
&= (X'X)^{-1}X'\text{Var}(y)X(X'X)^{-1} \\
&= (X'X)^{-1}X'\sigma^2 IX(X'X)^{-1} \\
&= (X'X)^{-1}\sigma^2
\end{aligned} \tag{6.3}
$$

and now lets referene (6.3) for god sake Note that the `echo = FALSE` parameter was added to the code chunk 6.1 to prevent printing of the R code that generated the plot.

Note that the `echo = FALSE` parameter was added to the code chunk 6.1 to prevent printing of the R code that generated the plot.

Note that the `echo = FALSE` parameter was added to the code chunk 6.1 to prevent printing of the R code that generated the plot. Note that the `echo = FALSE` parameter was added to the code chunk 6.1 to prevent printing of the R code that generated the plot.

6.3 level chapeter

fdsfdsf sdf fsd f sdf sd f sf

7 let's try an ew cahpeter

$$
\begin{aligned}
\text{Var}(\hat{\beta}) &= \text{Var}((X'X)^{-1}X'y) \\
&= (X'X)^{-1}X'\text{Var}(y)((X'X)^{-1}X')' \\
&= (X'X)^{-1}X'\text{Var}(y)X(X'X)^{-1} \\
&= (X'X)^{-1}X'\sigma^2 IX(X'X)^{-1} \\
&= (X'X)^{-1}\sigma^2
\end{aligned} \tag{7.1}
$$

trterter (7.1)

References

Bishop, Christopher M. 2006. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2001. *The Elements of Statistical Learning*. Springer Series in Statistics. New York, NY, USA: Springer New York Inc.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2014. *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated.