Linear Prediction

Deepu Vijayasenan

NITK, Surathkal

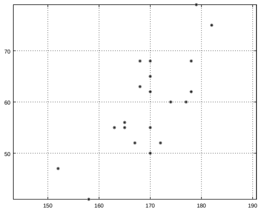
deepuv@nitk.ac.in

January 29, 2020

Learning

- Find a function that can learn the input output relation (*Model*)
- Determine the parameters of the model (*Learning*) from the available data (*Training data*)
- Simple example is learning a curve from a set of points
- What is a general function – Polynomials (Simplest one – Line)
- Supervised Learning Example – Linear Regression

$$y = mx + c$$



Linear Regression

- Model h(x) = m x + c
- Training Data

x	y
158	48
164	54
172	75
...	...

- Find *m* and *c* such that "total error" in predicting *y* (*Objective Function*) is minimum for *this dataset*

Linear Regression

- Model h(x) = m x + c
- Training Data

x	y
158	48
164	54
172	75
...	...

- Find m and c such that "total error" in predicting y (*Objective Function*) is minimum for *this dataset*

- Objective Function

$$
\begin{aligned}
E = & [48 - (m158 + c)]^2 \\
& + [54 - (m164 + c)]^2 \\
& + [75 - (m172 + c)]^2 + \dots
\end{aligned}
$$

Minimization of Objective Function

- To minimize we set $\frac{\partial E}{\partial m} = 0$ and $\frac{\partial E}{\partial c} = 0$

$$
\begin{aligned}
E &= \left[48 - (m158 + c)\right]^2 + \dots \\
\frac{\partial E}{\partial m} &= 2\left[48 - (m158 + c)\right](-158) + \dots &= 0 \\
\frac{\partial E}{\partial c} &= 2\left[48 - (m158 + c)\right](-1) + \dots &= 0
\end{aligned}
$$

- Two equations in two unknowns. Solving leads to $m \approx 3.06, c \approx -400$

Linear Regression

- Objective Function with one input and one output variable

$$
\begin{aligned}
E &= \sum_i [y_i - h(x_i)]^2 \\
E &= \sum_i [y_i - (mx_i + c)]^2 \\
\frac{\partial E}{\partial m} &= \sum_i 2\,[y_i - (mx_i + c)]\,(-x_i) && = 0 \\
\frac{\partial E}{\partial c} &= \sum_i 2\,[y_i - (mx_i + c)]\,(-1) && = 0 \\
\sum_i y_i x_i &= m \sum x_i^2 + c \sum_i x_i \\
\sum_i y_i &= m \sum x_i + c \sum_i 1
\end{aligned}
$$

Multiple variables

- Often multiple input variables are available
- Predict weight from height, age, shoulder size,...
- Training Data

$x^{(1)}$	$x^{(2)}$	y
158	18	48
164	35	54
172	24	75
...	...	...

- Extending the model

$$y = a_0 + a_1 x^{(1)} + a_2 x^{(2)} + \ldots + a_m x^{(m)}$$

- Modified Objective function

$$J = \sum_i \left[y_i - \left(a_0 + a_1 x^{(1)} + a_2 x^{(2)} \ldots + a_m x^{(m)} \right) \right]^2$$

Minimization

$$J = \sum_i \left[y_i - \left(a_0 + a_1 x^{(1)} + a_2 x^{(2)} \ldots + a_m x^{(m)} \right) \right]^2$$

$$\frac{\partial J}{\partial a_0} = \sum_i \left[y_i - \left(a_0 + a_1 x^{(1)} + a_2 x^{(2)} \ldots + a_m x^{(m)} \right) \right] (-1)$$

$$\frac{\partial J}{\partial a_l} = \sum_i \left[y_i - \left(a_0 + a_1 x^{(1)} + a_2 x^{(2)} \ldots + a_m x^{(m)} \right) \right] (-x^{(l)}), l = 1, \ldots, m$$

- (m+1) equations in (m+1) unknowns
- Solved to get the parameters

Linear Prediction

- Prediction of a signal sample from previous samples
- Predict $x[n]$ from past values $x[n-1], \ldots, x[n-P]$
- P^{th} order prediction

Formulation

$$x[n] = \sum_{i=1}^{P} a_i x[n-i]$$

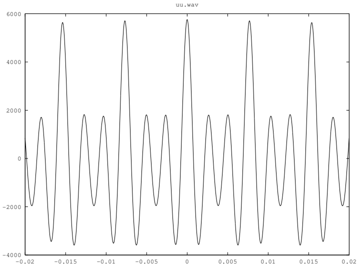
- P is typically 10-15
- Lot of signal available as training data
- estimate the same way as we did !!

Applications

- Remove the effect of glottal excitation
- Vocal tract modeling
- As features for ASR

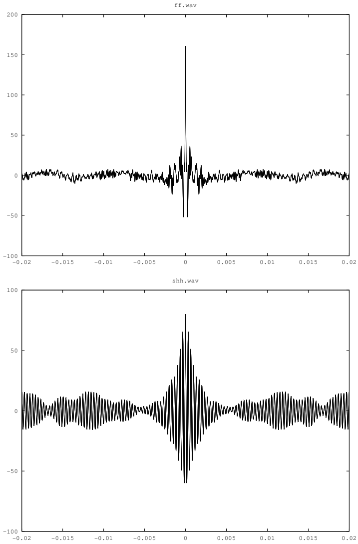
Autocorrelation

- Voiced signal will be periodic
- Autocorrelation will have peaks at fundamental frequency

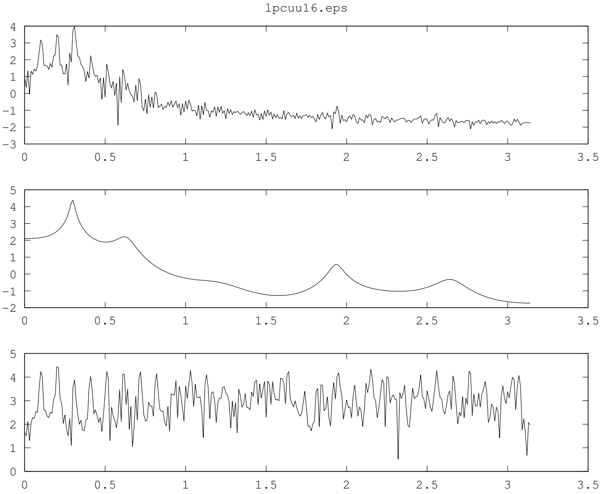


Autocorrelation

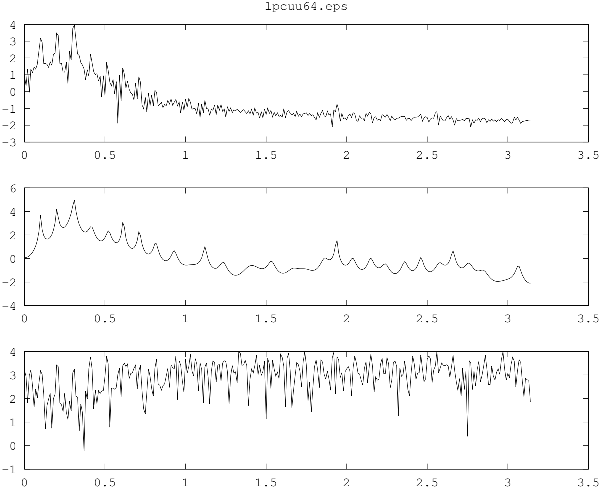
- Unvoiced signal will be periodic
- Assuming uncorrelated signal, autocorrelation is impulse like



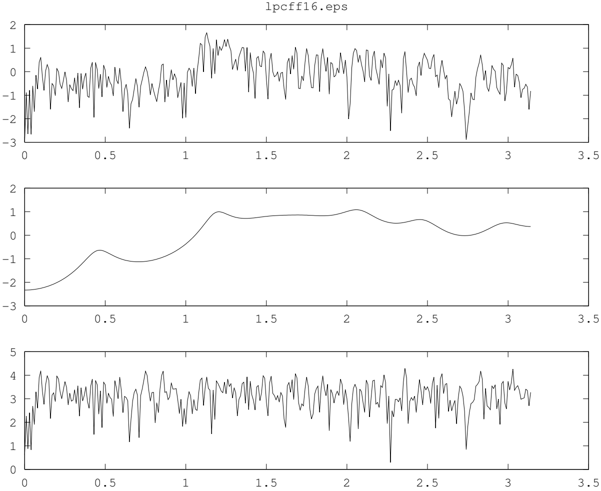
Linear Prediction of a Vowel



Linear Prediction of a Vowel



Linear Prediction of a fricative



Linear Prediction of a fricative