SMA 6304 / MIT 2.853 / MIT 2.854
Manufacturing Systems

Lecture 11: Forecasting

Lecturer: Prof. Duane S. Boning
Agenda

1. Regression
   • Polynomial regression
   • Example (using Excel)

2. Time Series Data & Regression
   • Autocorrelation – ACF
   • Example: white noise sequences
   • Example: autoregressive sequences
   • Example: moving average
   • ARIMA modeling and regression

3. Forecasting Examples
Regression – Review & Extensions

• Single Model Coefficient: Linear Dependence
  \[ \eta = \beta x \]

• Slope and Intercept (or Offset):
  \[ \eta = \beta_0 + \beta_1 x \]

• Polynomial and Higher Order Models:
  \[ \eta = \beta_0 + \beta_1 x + \beta_2 x^2 \]

• Multiple Parameters
  \[ \eta = \beta_0 + \beta_1 x + \beta_2 w \]

• Key point: “linear” regression can be used as long as the model is linear in the coefficients (doesn’t matter the dependence in the independent variable)
Polynomial Regression Example

- Replicate data provides opportunity to check for lack of fit

---

### Growth rate data

<table>
<thead>
<tr>
<th>observation number</th>
<th>amount of supplement (grams)</th>
<th>growth rate (coded units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>73</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>78</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>85</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>90</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>91</td>
</tr>
<tr>
<td>6</td>
<td>25</td>
<td>87</td>
</tr>
<tr>
<td>7</td>
<td>25</td>
<td>86</td>
</tr>
<tr>
<td>8</td>
<td>25</td>
<td>91</td>
</tr>
<tr>
<td>9</td>
<td>30</td>
<td>75</td>
</tr>
<tr>
<td>10</td>
<td>35</td>
<td>65</td>
</tr>
</tbody>
</table>
Growth Rate – First Order Model

• Mean significant, but linear term not
• Clear evidence of lack of fit

<table>
<thead>
<tr>
<th>source</th>
<th>sum of squares</th>
<th>degrees of freedom</th>
<th>mean square</th>
</tr>
</thead>
<tbody>
<tr>
<td>model</td>
<td>$S_M = 67,428.6$ mean 67,404.1 extra for linear 24.5</td>
<td>2 { 1 } 1</td>
<td>67,404.1</td>
</tr>
<tr>
<td>residual</td>
<td>$S_R = 686.4$ $S_L = 659.40$</td>
<td>8 { 4 } 4</td>
<td>85.8 164.85 ratio = 24.42</td>
</tr>
<tr>
<td>pure error</td>
<td>$S_E = 27.0$</td>
<td></td>
<td>6.75</td>
</tr>
<tr>
<td>total</td>
<td>$S_T = 68,115.0$</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>
Growth Rate – Second Order Model

- No evidence of lack of fit
- Quadratic term significant

### Analysis of variance for growth rate data: quadratic model

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of squares</th>
<th>Degrees of freedom</th>
<th>Mean square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>( S_M = 68,071.8 )</td>
<td>1</td>
<td>67,404.1</td>
</tr>
<tr>
<td></td>
<td>mean 67,404.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>extra for linear 24.5</td>
<td>3 ( \times ) 1</td>
<td>24.5</td>
</tr>
<tr>
<td></td>
<td>extra for quadratic 643.2</td>
<td>1</td>
<td>643.2</td>
</tr>
<tr>
<td>Residual</td>
<td>( S_R = 43.2 )</td>
<td>7 ( \times ) 3</td>
<td>5.40</td>
</tr>
<tr>
<td></td>
<td>( S_L = 16.2 )</td>
<td></td>
<td>ratio = 0.80</td>
</tr>
<tr>
<td></td>
<td>( S_E = 27.0 )</td>
<td>7 ( \times ) 4</td>
<td>6.75</td>
</tr>
<tr>
<td>Total</td>
<td>( S_T = 68,115.0 )</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>
Polynomial Regression In Excel

- Create additional input columns for each input
- Use “Data Analysis” and “Regression” tool

<table>
<thead>
<tr>
<th>x</th>
<th>x^2</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>100</td>
<td>73</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>78</td>
</tr>
<tr>
<td>15</td>
<td>225</td>
<td>85</td>
</tr>
<tr>
<td>20</td>
<td>400</td>
<td>90</td>
</tr>
<tr>
<td>20</td>
<td>400</td>
<td>91</td>
</tr>
<tr>
<td>25</td>
<td>625</td>
<td>87</td>
</tr>
<tr>
<td>25</td>
<td>625</td>
<td>86</td>
</tr>
<tr>
<td>25</td>
<td>625</td>
<td>91</td>
</tr>
<tr>
<td>30</td>
<td>900</td>
<td>75</td>
</tr>
<tr>
<td>35</td>
<td>1225</td>
<td>65</td>
</tr>
</tbody>
</table>

Regression Statistics

- Multiple R: 0.968
- R Square: 0.936
- Adjusted R Square: 0.918
- Standard Error: 2.541
- Observations: 10

ANOVA

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Significance F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>665.7</td>
<td>332.9</td>
<td>51.555</td>
<td>6.48E-05</td>
</tr>
<tr>
<td>Residual</td>
<td>7</td>
<td>45.19</td>
<td>6.456</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>710.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Coefficients Standard Error t Stat P-value Lower 95% Upper 95%
Intercept: 35.657 5.618 6.347 0.0004 22.373 48.942
x: 5.263 0.558 9.431 3.1E-05 3.943 6.582
x^2: -0.128 0.013 -9.966 2.2E-05 -0.158 -0.097
## Polynomial Regression

### Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Square</th>
<th>Mean Squar</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>665.70617</td>
<td>332.853</td>
<td>51.5551</td>
</tr>
<tr>
<td>Error</td>
<td>7</td>
<td>45.19383</td>
<td>6.456</td>
<td>Prob &gt; F</td>
</tr>
<tr>
<td>C. Total</td>
<td>9</td>
<td>710.90000</td>
<td></td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

### Lack Of Fit

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Square</th>
<th>Mean Squar</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack Of Fit</td>
<td>3</td>
<td>18.193829</td>
<td>6.0646</td>
<td>0.8985</td>
</tr>
<tr>
<td>Pure Error</td>
<td>4</td>
<td>27.000000</td>
<td>6.7500</td>
<td>Prob &gt; F</td>
</tr>
<tr>
<td>Total Error</td>
<td>7</td>
<td>45.193829</td>
<td></td>
<td>0.5157</td>
</tr>
</tbody>
</table>

### Summary of Fit

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RSquare</td>
<td>0.936427</td>
</tr>
<tr>
<td>RSquare Adj</td>
<td>0.918264</td>
</tr>
<tr>
<td>Root Mean Sq Error</td>
<td>2.540917</td>
</tr>
<tr>
<td>Mean of Response</td>
<td>82.1</td>
</tr>
<tr>
<td>Observations (or Sum Wgts)</td>
<td>10</td>
</tr>
</tbody>
</table>

### Parameter Estimates

| Term    | Estimate  | Std Error | t Ratio | Prob>|t| |
|---------|-----------|-----------|---------|-----|
| Intercept | 35.657437  | 5.617927  | 6.35    | 0.0004 |
|x         | 5.2628956  | 0.558022  | 9.43    | <.0001 |
x*x      | -0.127674  | 0.012811  | -9.97   | <.0001 |

### Effect Tests

<table>
<thead>
<tr>
<th>Source</th>
<th>Nparm</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>F Ratio</th>
<th>Prob &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>1</td>
<td>1</td>
<td>574.28553</td>
<td>88.9502</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
x*x    | 1     | 1  | 641.20451      | 99.3151 | <.0001   |

Copyright 2003 © Duane S. Boning.
Agenda

1. Regression
   • Polynomial regression
   • Example (using Excel)

2. Time Series Data & Time Series Regression
   • Autocorrelation – ACF
   • Example: white noise sequences
   • Example: autoregressive sequences
   • Example: moving average
   • ARIMA modeling and regression

3. Forecasting Examples
Time Series – Time as an Implicit Parameter

- Data is often collected with a time-order.
- An underlying dynamic process (e.g. due to physics of a manufacturing process) may create autocorrelation in the data.
Intuition: Where Does Autocorrelation Come From?

- Consider a chamber with volume $V$, and with gas flow in and gas flow out at rate $f$. We are interested in the concentration $x$ at the output, in relation to a known input concentration $w$.

\[
\frac{dx_t}{dt} = (w_t - x_t) \frac{f}{V}
\]

\[
x_t = w_t - \frac{V}{f} \frac{dx_t}{dt} = w_t - \tau \frac{dx_t}{dt}
\]

Consider a step change in input of $w_0$ at $t = 0$. Then

\[
x_t = w_0(1 - e^{-t/\tau})
\]

Discretizing:

\[
x_t = x_{t-1} + (w_0 - x_{t-1})(1 - e^{-\Delta t/T})
\]

\[
x_t = aw_t + (1 - a)x_{t-1} \quad \text{where} \quad a = 1 - e^{-\Delta t/T}
\]

The correlation between $x_t$ & $x_{t-1}$ is $\rho = 1 - a = e^{-\Delta t/T}$
**Key Tool: Autocorrelation Function (ACF)**

- Time series data: time index \( i \)
  \[ x_i \sim \mathcal{N}(0, 1) \]

- CCF: cross-correlation function
  \[ r_{xy}(k) = \frac{1}{N} \sum_{i=1}^{N-1} \frac{(x_i - \bar{x})(y_{i+k} - \bar{y})}{s_xs_y} \]

- ACF: auto-correlation function
  \[ r_{xx}(k) = \frac{1}{N} \sum_{i=1}^{N-1} \frac{(x_i - \bar{x})(x_{i+k} - \bar{x})}{s_x^2} \]

⇒ ACF shows the “similarity” of a signal to a lagged version of same signal
Stationary vs. Non-Stationary

Stationary series: Process has a **fixed** mean
White Noise – An Uncorrelated Series

- Data drawn from IID gaussian
  \[ w_i \sim N(0, 1) \]

- ACF: We also plot the $3\sigma$ limits – values within these not significant

- Note that $r(0) = 1$ always (a signal is always equal to itself with zero lag – perfectly autocorrelated at $k = 0$)

- Sample mean
  \[ \bar{w} = \frac{1}{N} \sum_{i}^{N} w_i \]

- Sample variance
  \[ s_w^2 = \frac{1}{N - 1} \sum_{i}^{N} (w_i - \bar{w})^2 \]
Autoregressive Disturbances

- Generated by:
  \[ w_i \sim N(0, 1) \]
  \[ c_i = \alpha \cdot c_{i-1} + w_i \]
  Shown: \( \alpha = 0.9 \)

- Mean
  \[ \mu_c = E(c_i) = 0 \]
  since \( \mu_w = 0 \)

- Variance
  \[ \sigma^2_c = \text{Var}(c_i) = E[(c_i - \bar{c})^2] \]
  \[ = E(\alpha^2 c_{i-1}^2 + 2\alpha c_{i-1}w_i + w_i^2) \]
  \[ = \alpha^2 \text{Var}(c_i) + \text{Var}(w_i) \]

\[ \Rightarrow \sigma^2_c = \frac{\sigma^2_w}{1 - \alpha^2} \]

So AR (autoregressive) behavior increases variance of signal.

Copyright 2003 © Duane S. Boning.
Another Autoregressive Series

• Generated by:
  \[ w_i \sim N(0, 1) \]
  \[ c_i = \alpha \cdot c_{i-1} + w_i \]
  Shown: \( \alpha = -0.9 \)

• High negative autocorrelation:

Slow drop in ACF with large \( \alpha \)

But now ACF alternates in sign
Random Walk Disturbances

• Generated by:
  
  \[ w_i \sim N(0, 1) \]
  
  \[ c_i = 1 \cdot c_{i-1} + w_i \]
  
  AR with \( \alpha = 1 \)

• Mean
  
  \( \bar{c} \neq 0 \) non-stationary

• Variance
  
  Variance increases as sequence gets longer

Very slow drop in ACF for \( \alpha = 1 \)
Moving Average Sequence

• Generated by:
  \[ w_i \sim N(0, 1) \]
  \[ c_i = w_i + \beta \cdot w_{i-1} \]
  Shown: \( \beta = 0.5 \)

• Mean
  \[ \mu_c = E(c_i) = 0 \]
  since \( \mu_w = 0 \)

• Variance
  \[ \sigma_c^2 = \text{Var}(c_i) = E[(c_i - \bar{c})^2] \]
  \[ = E(w_i^2 + 2\beta w_i w_{i-1} + \beta^2 w_{i-1}^2) \]
  \[ = (1 + \beta^2) \text{Var}(w_i) \]

\[ \Rightarrow \sigma_c^2 = (1 + \beta^2) \sigma_w^2 \]

So MA (moving average) behavior also increases variance of signal.

Copyright 2003 © Duane S. Boning.
ARMA Sequence

- Generated by:
  \[ w_i \sim N(0, 1) \]
  \[ c_i = \alpha \cdot c_{i-1} + w_i + \beta \cdot w_{i-1} \]
  Shown: \( \alpha = 0.9, \beta = 0.5 \)

- Both AR & MA behavior

Slow drop in ACF with large \( \alpha \)
ARIMA Sequence

- Start with ARMA sequence:
  \[ w_i \sim N(0, 1) \]
  \[ c_i = \alpha \cdot c_{i-1} + w_i + \beta \cdot w_{i-1} \]
  Shown: \( \alpha = 0.9, \quad \beta = 0.5 \)

- Add Integrated (I) behavior
  \[ x_i = x_{i-1} + c_i \]

random walk (integrative) action

Slow drop in ACF with large \( \alpha \)
Periodic Signal with Autoregressive Noise

Original Signal

After Differencing

\[ d_i = x_i - x_{i-1} \]

See underlying signal with period = 5
Agenda

1. Regression
   • Polynomial regression
   • Example (using Excel)

2. Time Series Data & Regression
   • Autocorrelation – ACF
   • Example: white noise sequences
   • Example: autoregressive sequences
   • Example: moving average
   • ARIMA modeling and regression

3. Forecasting Examples
Cross-Correlation: A Leading Indicator

• Now we have two series:
  – An “input” or explanatory variable $x$
  – An “output” variable $y$

$$y_i = x_{i-k} + w_i$$

$$w_i \sim N(0, 1)$$

Shown: lag $k = 20$ and autoregressive $x$ with $\alpha = 0.9$

• CCF indicates both AR and lag:
The ACF or CCF are helpful tools in selecting an appropriate model structure
  – Autoregressive terms?
    • $x_i = \alpha x_{i-1}$
  – Lag terms?
    • $y_i = \gamma x_{i-k}$

One can structure data and perform regressions
  – Estimate model coefficient values, significance, and confidence intervals
  – Determine confidence intervals on output
  – Check residuals
1. Statistical Fundamentals
   • Sampling distributions
   • Point and interval estimation
   • Hypothesis testing

2. Regression
   • ANOVA
   • Nominal data: modeling of treatment effects (mean differences)
   • Continuous data: least square regression $y = f(x, b)$

3. Time Series Data & Forecasting
   • Autoregressive, moving average, and integrative behavior
   • Auto- and Cross-correlation functions
   • Regression and time-series modeling
     $x_i = f(x_i, b)$
     $y_i = f(x_i, b)$