Linear Regression Models
P8111

Lecture 12

Jeff Goldsmith
March 1, 2016

Today's Lecture

- Gauss-Markov theorem
- Maximum likelihood inference
- Regression diagnostics

What's so great about LSEs?

- Nice projection-space interpretation
- They're the "best" linear unbiased estimators
- They're maximum likelihood estimators under Normally-distributed errors

Gauss-Markov theorem

Assume the model

$$y = X\beta + \epsilon$$

where $E(\epsilon) = 0$ and $Var(\epsilon) = \sigma^2 I$. Also assume X is a full rank design matrix.

- Among all unbiased linear estimators Cy of the regression coefficients β, the LSE has minimum variance and is unique.

We call the LSEs "BLUE".

Gauss-Markov theorem – proof

Gauss-Markov theorem – proof

Gauss-Markov theorem – caveats

The Gauss-Markov theorem is great, but notice the details:

- Assumed $Var(\epsilon) = \sigma^2 I$
- Only talking about unbiased linear estimators

Maximum likelihood estimation

Continue assuming the model

$$y = X\beta + \epsilon$$

where $E(\epsilon) = 0$ and $Var(\epsilon) = \sigma^2 I$.

- Additionally, assume $\epsilon \sim N(0, \sigma^2 I)$
- Put differently, we're imposing the model

$$y \sim N(X\beta, \sigma^2 I)$$

- y is multivariate Normal with uncorrelated entries; the y_i are each independently Normally distributed

Maximum likelihood estimation

Using independently Normal y_i's:

Maximum likelihood estimation

Using matrix notation:

Regression diagnostics

- Regression diagnostics are tools used to determine whether a given model is consistent with the data
- Usually focus on residuals
- Recall that fitted values are given by $\hat{y} = Hy$ where H is the hat matrix
- Residuals are defined as $y - \hat{y} = (I - H)y$

$\hat{\epsilon}$ and ϵ

- $E(\hat{\epsilon}) =$

- $Var(\hat{\epsilon}) =$

- Residuals are mean zero, but don't have constant variance nor are they uncorrelated.

$\hat{\epsilon}$ and ϵ

- $\hat{\epsilon} = \qquad\qquad\qquad = (I - H)\epsilon$

- If ϵ is Normally distributed, so are the residuals
- Also note residuals sum to zero

Residuals and fitted values

$$
\begin{aligned}
Cov(\hat{\boldsymbol{\epsilon}}^T, \hat{\boldsymbol{y}}) &= Cov((I - \boldsymbol{H})\boldsymbol{y}, \boldsymbol{H}\boldsymbol{y}) \\
&= (I - \boldsymbol{H})\sigma^2 I \boldsymbol{H}^T \\
&= \sigma^2 (\boldsymbol{H} - \boldsymbol{H})
\end{aligned}
$$
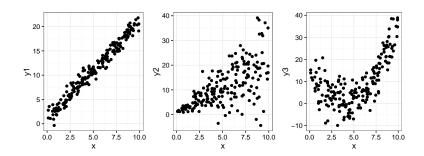
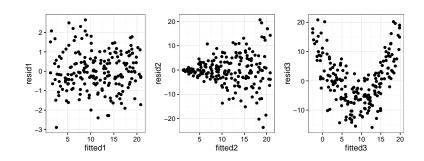- So residuals and fitted values are uncorrelated

Residuals when model is correct

- Often we plot the residuals against one of the predictors or against the fitted values
- What we look for:
  - $E(\hat{\epsilon}|x) = 0$
  - $V(\hat{\epsilon}|x) = \sigma^2(1 - h_{ii})$
- If the model is incorrect, you may be able to spot:
  - Patterns in the residuals
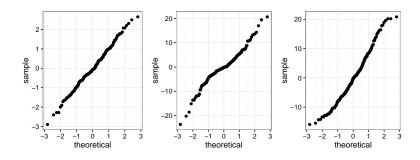 - Clear non-constant variance

Some data plots

Some residual plots

Checking Normality assumption

- We often assume Normality for the errors
- Useful to check Normality of residuals
- Try a QQ plot:
  - Compute the sample quantiles of the residuals
  - Compute the quantiles of a standard Normal of size n
  - Plot these against each other
- Can also use the Shapiro-Wilk test based on correlation between observed and theoretical quantiles

Checking Normality assumption

Checking model structure

- You can plot residuals against each of the predictors, or plot outcomes against predictors
- Keep in mind the MLR uses adjusted relationships; scatterplots don't show that adjustment
- Adjusted variable plots (partial regression plots, added variable plots) can be useful

Adjusted variable plots

- Regress y on everything but x_j; take residuals $r_{y|-x_j}$
- Regress x_j on everything but x_j; take residuals $r_{x_j|-x_j}$
- Regress $r_{y|-x_j}$ on $r_{x_j|-x_j}$; slope of this line will match β_j in the full MLR
- Plot of $r_{y|-x_j}$ against $r_{x_j|-x_j}$ shows the "adjusted" relationship

What should you do ...

if your assumptions are violated?

- Depends on the assumption
- For problems with the errors, use LSE anyway; maybe use bootstrap for inference
- For non-linearity, try an augmented model

Today's big ideas

- Gauss-Markov, MLE, regression diagnostics

- Suggested reading: Faraway Ch 2.8, Ch 7