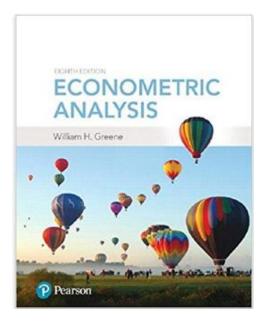
# **Econometrics** I

Professor William Greene Stern School of Business Department of Economics



# **Econometrics** I

#### Part 3 – Least Squares Algebra

# Vocabulary

- **Some terms** to be used in the discussion.
  - Population characteristics and entities vs. sample quantities and analogs
  - Residuals and disturbances
  - Population regression line and sample regression
- **Objective**: Learn about the conditional mean function. 'Estimate' β and  $σ^2$
- First step: Mechanics of fitting a line (hyperplane) to a set of data

# **Fitting Criteria**

- The set of points in the sample
- Fitting criteria what are they:
  - LAD: Minimize<sub>b</sub>  $\Sigma |y x'b_{LAD}|$
  - Least squares: Minimize<sub>b</sub>  $\Sigma$  (y **x'b<sub>LS</sub>**)<sup>2</sup>
  - and so on
- Why least squares?

A fundamental result:

# Sample moments are "good" estimators of their population counterparts

We will examine this principle and apply it to least squares computation.

# An Analogy Principle for Estimating β

In the population

Continuing (assumed) Summing, Exchange  $\Sigma_i$  and E[]

$$E[\mathbf{y} \mid \mathbf{X}] = \mathbf{X}\beta \text{ so}$$

$$E[\mathbf{y} - \mathbf{X}\beta \mid \mathbf{X}] = \mathbf{0}$$

$$E[\mathbf{x}_i \varepsilon_i] = \mathbf{0} \text{ for every } i$$

$$\sum_i E[\mathbf{x}_i \varepsilon_i] = \sum_i \mathbf{0} = \mathbf{0}$$

$$E[\sum_i \mathbf{x}_i \varepsilon_i] = E[\mathbf{X}'\varepsilon] = \mathbf{0}$$

$$E[\mathbf{X}'(\mathbf{y} - \mathbf{X}\beta)] = \mathbf{0}$$

So, if  $\mathbf{X}\beta$  is the conditional mean, then  $E[\mathbf{X}'\epsilon] = \mathbf{0}$ . We choose **b**, the estimator of  $\beta$ , to mimic this population result: i.e., mimic the population mean with the sample mean

Find **b** such that 
$$\frac{1}{n}\mathbf{X'e} = \mathbf{0} = \frac{1}{n}\mathbf{X'}(\mathbf{y} - \mathbf{Xb})$$

As we will see, the solution is the least squares coefficient vector.

#### **Population Moments**

We assumed that  $E[\varepsilon_i | \mathbf{x}_i] = 0$ . (Slide 2:40) It follows that  $Cov[\mathbf{x}_i, \varepsilon_i] = \mathbf{0}$ . Proof:  $Cov(\mathbf{x}_i, \varepsilon_i) = Cov(\mathbf{x}_i, E[\varepsilon_i | \mathbf{x}_i]) = Cov(\mathbf{x}_i, 0) = \mathbf{0}$ . (Theorem B.2). If  $E[y_i | \mathbf{x}_i] = \mathbf{x}_i \,\beta$ , then  $\beta = (Var[\mathbf{x}_i])^{-1} Cov[\mathbf{x}_i, y_i]$ . Proof:  $Cov[\mathbf{x}_i, y_i] = Cov[\mathbf{x}_i, E[y_i | \mathbf{x}_i]] = Cov[\mathbf{x}_i, \mathbf{x}_i \,\beta]$ This will provide a population analog to the statistics we compute with the data.

### U.S. Gasoline Market, 1960-1995

G	CONST	PG	Y
129.7	1	0.925	6036
131.3	1	0.914	6113
137.1	1	0.919	6271
141.6	1	0.918	6378
148.8	1	0.914	6727
155.9	1	0.949	7027
164.9	1	0.97	7280
171	1	1	7513
183.4	1	1.014	7728
195.8	1	1.047	7891
207.4	1	1.056	8134
218.3	1	1.063	8322
226.8	1	1.076	8562
237.9	1	1.181	9042
225.8	1	1.599	8867
232.4	1	1.708	8944
241.7	1	1.779	9175
249.2	1	1.882	9381
261.3	1	1.963	9735
248.9	1	2.656	9829
226.8	1	3.691	9722
225.6	1	4.109	9769
228.8	1	3.894	9725
239.6	1	3.764	9930
244.7	1	3.707	10421
245.8	1	3.738	10563
269.4	1	2.921	10780
276.8	1	3.038	10859
279.9	1	3.065	11186
284.1	1	3.353	11300
282	1	3.834	11389
271.8	1	3.766	11272
280.2	1	3.751	11466
286.7	1	3.713	11476
290.2	1	3.732	11636
297.8	1	3.789	11934

3-7/29

# **Least Squares**

#### ■ Example will be, $G_i$ regressed on $\mathbf{x}_i = [1, PG_i, Y_i]$

■ Fitting criterion: Fitted equation will be  $y_i = b_1 x_{i1} + b_2 x_{i2} + ... + b_K x_{iK}.$ 

Criterion is based on residuals:

 $e_i = y_i - b_1 x_{i1} + b_2 x_{i2} + ... + b_K x_{iK}$ Make  $e_i$  as small as possible. Form a criterion and minimize it.

# **Fitting Criteria**

- **Sum of residuals:**  $\sum_{i=1}^{n} e_i$
- **u** Sum of squares:  $\sum_{i=1}^{n} e_i^2$
- **D** Sum of absolute values of residuals:  $\sum_{i=1}^{n} |e_i|$
- □ Absolute value of sum of residuals  $\left|\sum_{i=1}^{n} e_{i}\right|$
- We focus on  $\sum_{i=1}^{n} e_i^2$  now and  $\sum_{i=1}^{n} |e_i|$  later

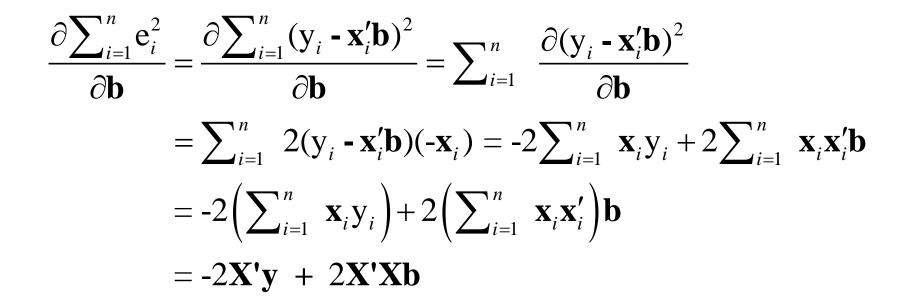
### Least Squares Algebra

$$\sum_{i=1}^{n} \mathbf{e}_{i}^{2} = \sum_{i=1}^{n} (\mathbf{y}_{i} - \mathbf{x}_{i}'\mathbf{b})^{2} = \mathbf{e}'\mathbf{e} = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b})$$

Matrix and vector derivatives.

Derivative of a scalar with respect to a vector Derivative of a column vector wrt a row vector Other derivatives

#### Least Squares Normal Equations



3-11/29

#### Least Squares Normal Equations

$$\frac{\partial (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b})}{\partial \mathbf{b}} = -2\mathbf{X}'(\mathbf{y} - \mathbf{X}\mathbf{b}) = \mathbf{0}$$
$$\frac{\partial (1 \times 1)}{\partial (K \times 1)} \quad (-2)(n \times K)'(n \times 1)$$
$$= (-2)(K \times n)(n \times 1) = K \times 1$$

Note: Derivative of  $(1 \times 1)$  wrt K  $\times 1$  vector is a K  $\times 1$  vector.

Solution:  $-2\mathbf{X'}(\mathbf{y} - \mathbf{X}\mathbf{b}) = \mathbf{0} \Rightarrow \mathbf{X'}\mathbf{y} = \mathbf{X'}\mathbf{X}\mathbf{b}$ 

#### **Least Squares Solution**

Assuming it exists:  $\mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'y}$ Note the analogy:  $\boldsymbol{\beta} = (\operatorname{Var}(\mathbf{x}))^{-1} (\operatorname{Cov}(\mathbf{x},\mathbf{y}))$  $\mathbf{b} = \left(\frac{1}{n}\mathbf{X'X}\right)^{-1} \left(\frac{1}{n}\mathbf{X'y}\right) = \left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{i}\mathbf{x}_{i}'\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{i}y_{i}\right)$ 

Suggests something desirable about least squares

3-13/29

#### **Second Order Conditions**

**Necessary Condition :** First derivatives =  $\mathbf{0}$  $\partial (\mathbf{v} - \mathbf{X}\mathbf{b})'(\mathbf{v} - \mathbf{X}\mathbf{b})$ 

$$\frac{\partial (\mathbf{y} - \mathbf{X}\mathbf{b}) (\mathbf{y} - \mathbf{X}\mathbf{b})}{\partial \mathbf{b}} = -2\mathbf{X'}(\mathbf{y} - \mathbf{X}\mathbf{b})$$

Sufficient Condition : Second derivatives ...

$$\frac{\partial^{2}(\mathbf{y} \cdot \mathbf{X}\mathbf{b})'(\mathbf{y} \cdot \mathbf{X}\mathbf{b})}{\partial \mathbf{b} \partial \mathbf{b}'} = \frac{\partial \left(\frac{\partial (\mathbf{y} \cdot \mathbf{X}\mathbf{b})'(\mathbf{y} \cdot \mathbf{X}\mathbf{b})}{\partial \mathbf{b}'}\right)}{\partial \mathbf{b}'}$$
$$= \frac{\partial \left\{-2\mathbf{X}'(\mathbf{y} \cdot \mathbf{X}\mathbf{b})\right\}}{\partial \mathbf{b}'} = \frac{\partial \left(-2\mathbf{X}'\mathbf{y}\right)}{\partial \mathbf{b}'} + \frac{\partial \left\{-2\mathbf{X}'(\mathbf{-X}\mathbf{b})\right\}}{\partial \mathbf{b}'} = \mathbf{0} + \frac{\partial 2\mathbf{X}'\mathbf{X}\mathbf{b}}{\partial \mathbf{b}'}$$
$$= \frac{\partial \mathbf{K} \times 1 \text{ column vector}}{\partial \mathbf{1} \times \mathbf{K} \text{ row vector}} = \mathbf{K} \times \mathbf{K} \text{ matrix}$$
$$= 2\mathbf{X}'\mathbf{X}$$

# Side Result: Sample Moments

$$\mathbf{X'X} = \begin{bmatrix} \sum_{i=1}^{n} x_{i1}^{2} & \sum_{i=1}^{n} x_{i1} x_{i2} & \dots & \sum_{i=1}^{n} x_{i1} x_{iK} \\ \sum_{i=1}^{n} x_{i2} x_{i1} & \sum_{i=1}^{n} x_{i2}^{2} & \dots & \sum_{i=1}^{n} x_{i2} x_{iK} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{n} x_{iK} x_{i1} & \sum_{i=1}^{n} x_{iK} x_{i2} & \dots & \sum_{i=1}^{n} x_{iK}^{2} \end{bmatrix}$$
$$= \sum_{i=1}^{n} \begin{bmatrix} x_{i1}^{2} & x_{i1} x_{i2} & \dots & x_{i1} x_{iK} \\ x_{i2} x_{i1} & x_{i2}^{2} & \dots & x_{i2} x_{iK} \\ \dots & \dots & \dots & \dots \\ x_{iK} x_{i1} & x_{iK} x_{i2} & \dots & x_{iK}^{2} \end{bmatrix}$$
$$= \sum_{i=1}^{n} \begin{bmatrix} x_{i1} \\ x_{i2} \\ \dots \\ x_{iK} \end{bmatrix} [x_{i1} & x_{i2} & \dots & x_{iK}]$$
$$= \sum_{i=1}^{n} \mathbf{x}_{i} \mathbf{x}_{i}'$$

3-15/29

Part 3: Least Squares Algebra

#### Does **b** Minimize **e'e**?

$$\frac{\partial^{2} \mathbf{e' e}}{\partial \mathbf{b} \partial \mathbf{b'}} = 2\mathbf{X' X} = 2 \begin{bmatrix} \sum_{i=1}^{n} x_{i1}^{2} & \sum_{i=1}^{n} x_{i1} x_{i2} & \dots & \sum_{i=1}^{n} x_{i1} x_{iK} \\ \sum_{i=1}^{n} x_{i2} x_{i1} & \sum_{i=1}^{n} x_{i2}^{2} & \dots & \sum_{i=1}^{n} x_{i2} x_{iK} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{n} x_{iK} x_{i1} & \sum_{i=1}^{n} x_{iK} x_{i2} & \dots & \sum_{i=1}^{n} x_{iK}^{2} \end{bmatrix}$$

If there were a single b, we would require this to be positive, which it would be;  $2\mathbf{x'x} = 2\sum_{i=1}^{n} x_i^2 > 0$ . OK The matrix counterpart of a positive number is a **positive definite matrix**.

3-16/29

# **A Positive Definite Matrix**

Matrix C is positive definite if a'Ca is > 0 for every a.
Generally hard to check. Requires a look at characteristic roots (later in the course).
For some matrices, it is easy to verify. X'X is one of these.

**a'X'Xa** = (**a'X'**)(**Xa**) = (**Xa**)'(**Xa**) = **v'v** =  $\sum_{k=1}^{K} v_k^2 > 0$ Could **v** = **0**? **v** = **0** means **Xa** = **0**. Is this possible? No. Conclusion: **b** = (**X'X**)<sup>-1</sup>**X'y** does indeed minimize **e'e**.

# Algebraic Results - 1

In the population :  $E[X'\epsilon] = 0$ 

- In the sample:  $\frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_{i}\mathbf{e}_{i} = \mathbf{0}$
- $\mathbf{X}'\mathbf{e} = \mathbf{0}$  means for each column of  $\mathbf{X}$ ,  $\mathbf{x}'_k \mathbf{e} = 0$

(1) Each column of  $\mathbf{X}$  is orthogonal to  $\mathbf{e}$ .

(2) One of the columns of  $\mathbf{X}$  is a column of ones.

**i'e** = 
$$\sum_{i=1}^{n} e_i = 0$$
. The residuals sum to zero.  
(3) It follows that  $\frac{1}{n} \sum_{i=1}^{n} e_i = 0$  which mimics  $E[\varepsilon_i] = 0$ .

# Residuals vs. Disturbances

Disturbances (population)  $y_i - \mathbf{x}'_i \boldsymbol{\beta} = \varepsilon_i$ Partitioning **y**:  $\mathbf{y} = E[\mathbf{y}|\mathbf{X}] + \varepsilon$ = conditional mean + disturbance Residuals (sample)  $y_i - \mathbf{x}'_i \mathbf{b} = e_i$ Partitioning **y**:  $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$ = projection + residual

Note : Projection into the column space of X, i.e., the set of linear combinations of the columns of X. Xb is one of these.)

# Algebraic Results - 2

- $\square A "residual maker" \mathbf{M} = (\mathbf{I} \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'})$
- $\Box e = y Xb = y X(X'X)^{-1}X'y = My$
- **My** = The residuals that result when **y** is regressed on **X**
- MX = 0 (This result is fundamental!)
  - How do we interpret this result in terms of residuals? When a column of **X** is regressed on all of **X**, we get a perfect fit and zero residuals.
- (Therefore) My = MXb + Me = Me = e
   (You should be able to prove this.

□ 
$$y = Py + My, P = X(X'X)^{-1}X' = (I - M).$$
  
PM = MP = 0.

**Py** is the projection of **y** into the column space of **X**.

# The M Matrix

- $\square M = I X(X'X)^{-1}X' \text{ is an nxn matrix}$
- M is <u>symmetric</u> M = M'
- M is <u>idempotent</u> M\*M = M

(just multiply it out)

■ M is singular; M<sup>-1</sup> does not exist.

(We will prove this later as a side result in another derivation.)

# Results when X Contains a Constant Term

 $y = \mathbf{x}'\mathbf{b}$  (the regression line passes through the means) These do not apply if the model has no constant term.

### U.S. Gasoline Market, 1960-1995

G	CONST	PG	Y
129.7	1	0.925	6036
131.3	1	0.914	6113
137.1	1	0.919	6271
141.6	1	0.918	6378
148.8	1	0.914	6727
155.9	1	0.949	7027
164.9	1	0.97	7280
171	1	1	7513
183.4	1	1.014	7728
195.8	1	1.047	7891
207.4	1	1.056	8134
218.3	1	1.063	8322
226.8	1	1.076	8562
237.9	1	1.181	9042
225.8	1	1.599	8867
232.4	1	1.708	8944
241.7	1	1.779	9175
249.2	1	1.882	9381
261.3	1	1.963	9735
248.9	1	2.656	9829
226.8	1	3.691	9722
225.6	1	4.109	9769
228.8	1	3.894	9725
239.6	1	3.764	9930
244.7	1	3.707	10421
245.8	1	3.738	10563
269.4	1	2.921	10780
276.8	1	3.038	10859
279.9	1	3.065	11186
284.1	1	3.353	11300
282	1	3.834	11389
271.8	1	3.766	11272
280.2	1	3.751	11466
286.7	1	3.713	11476
290.2	1	3.732	11636
297.8	1	3.789	11934

3-23/29

#### Least Squares Algebra

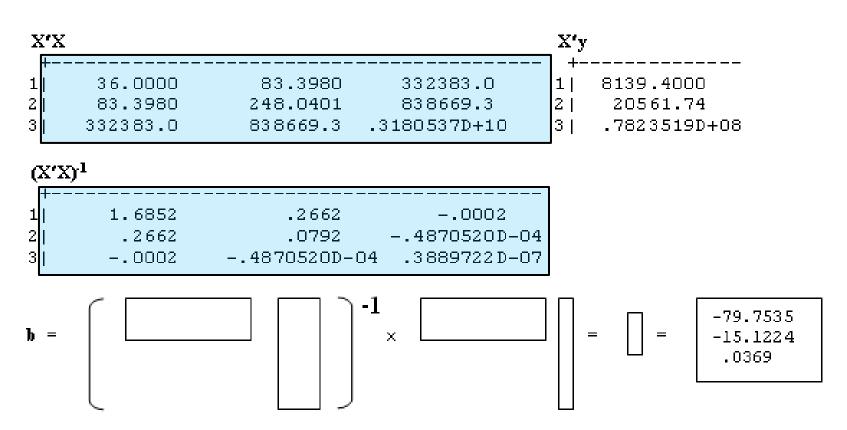
129.7		1	0.925	6036
131.3		1	0.914	6113
137.1		1	0.919	6271
141.6		1	0.918	6378
148.8		1	0.914	6727
155.9		1	0.949	7027
164.9		1	0.97	7280
171		1	1	7513
183.4		1	1.014	7728
195.8		1	1.047	7891
207.4		1	1.056	8134
218.3		1	1.063	8322
226.8		1	1.076	8562
237.9		1	1.181	9042
225.8		1	1.599	8867
232.4		1	1.708	8944
241.7		1	1.779	9175
249.2		1	1.882	9381
261.3		1	1.963	9735
248.9		1	2.656	9829
226.8	X =	1	3.691	9722
225.6		1	4.109	9769
228.8		1	3.894	9725
239.6		1	3.764	9930
244.7		1	3.707	10421
245.8		1	3.738	10563
269.4		1	2.921	10780
276.8		1	3.038	10859
279.9		1	3.065	11186
284.1		1	3.353	11300
282		1	3.834	11389
271.8		1	3.766	11272
280.2		1	3.751	11466
286.7		1	3.713	11476
290.2		1	3.732	11636
297.8		1	3.789	11934

y =

3-24/29

#### Part 3: Least Squares Algebra

#### Least Squares



Re			
	UIU	<b>u</b>	

129.7	
131.3	
137.1	
141.6	
155.9	
141.6 148.8 155.9 164.9 171 183.4 195.8 207.4 218.3 225.9	
171	
183.4	
195.8	
218.3	
226.8	
226.8 237.9	
225.8 232.4	
232.4	
241.7 249.2	
261.3	
261.3 248.9	
226.8	
225.6	
228.8 239.6	
244.7	
245.8	
269.4	
276.8	
279.9 284.1	
282	
271.8	
280.2	
286.7 290.2	
290.2 297.8	

			l	•	
	0.925 0.914 0.919 0.918 0.914 0.949 0.97 1 1.014 1.047 1.056 1.063 1.076 1.181 1.599 1.708 1.708 1.709 1.882 1.963 2.656 3.691 4.109 3.894 3.764 3.707 3.738 2.921 3.038 3.065	6036 6113 6271 6378 6727 7027 7280 7513 7728 7891 8134 8322 8562 9042 8867 8944 9175 9381 9735 9829 9722 9769 9725 9930 10421 10563 10780 10859 11186	-79.7535 -15.1224 .0369	=	0.590391 -0.818823 -0.776628 -0.242229 -5.98792 -9.43475 -9.45803 -11.5068 -6.83297 0.0480476 2.8125 6.87733 6.71303 1.67911 2.36131 7.76678 9.61187 11.0639 11.319 5.92825 3.43036 6.81625 8.38944 9.65486 -4.23501 -7.90891 -4.67559 1.57702 -6.98764
1 1	3.707 3.738	10421 10563			-4.23501 -7.90891
1	3.038	10859			1.57702 -6.98764
1 1	3. 834 3. 766 3. 751	11300 11389 11272 11466			-2.64132 -0.753384 -7.66202
1 1 1	3, 751 3, 713 3, 732 3, 789	11466 11476 11636 11934			-6.65141 -1.09526 -3.21519 -5.75549

e =

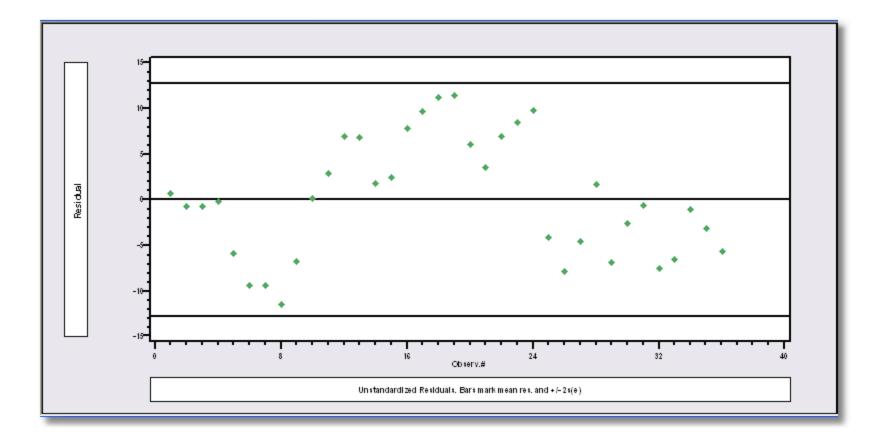
П

-

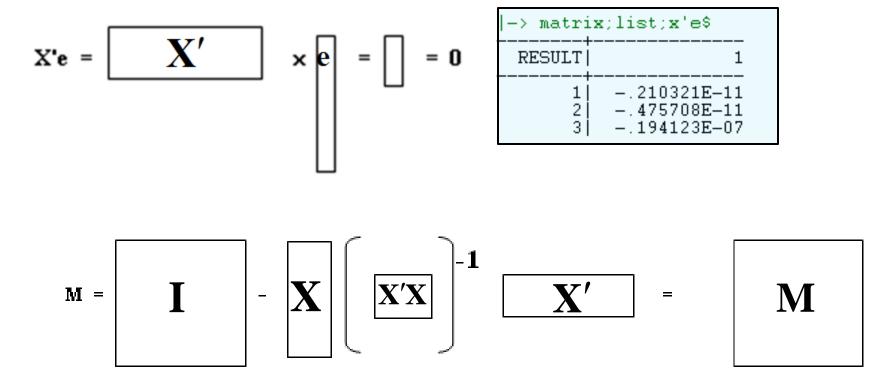
#### 3-26/29

Part 3: Least Squares Algebra

### Least Squares Residuals (autocorrelated)



#### Least Squares Algebra-3



M is  $n \times n$  potentially huge

3-28/29

#### Least Squares Algebra-4

[36, 3]	Cell: 3,4015	1e-010	✓ ×	
	1	2	3	
1	2.35922e-014	6.06459e-014	2.09639e-010	
2	3.35287e-014	8.32667e-014	3.00815e-010	
3	7.77156e-015	2.04559e-014	6.27551e-011	
4	2.10942e-015	5.96745e-015	7.61702e-012	
5	3.63043e-014	8.9373e-014	3.34808e-010	
6	1.12688e-014	2.65066e-014	9.98739e-011	
7	2.83662e-014	6.73905e-014	2.6381e-010	
8	2.68535e-014	6.53089e-014	2.54417e-010	
9	2.14967e-014	5.13235e-014	2.04778e-010	
10	2.25098e-014	5.48901e-014	2.15817e-010	
11	8.17402e-015	1.86101e-014	8.21814e-011	
12	1.18239e-014	2.77348e-014	1.20366e-010	
13	3.76088e-014	8.98864e-014	3.64821e-010	
14	3.41394e-014	8.24063e-014	3.3674e-010	
15	3.31124e-014	8.01442e-014	3.18664e-010	
16	2.38976e-014	5.67046e-014	2.29704e-010	
17	3.7248e-014	8.75133e-014	3.53964e-010	
18	7.16094e-015	1.31006e-014	6.99174e-011	
19	1.77636e-014	4.14668e-014	1.7792e-010	
20	2.55143e-014	6.09235e-014	2.41414e-010	
21	2.65898e-014	7.01522e-014	2.37378e-010	
22	2.36478e-014	6.21864e-014	2.04267e-010	
23	3.65263e-014	9.32449e-014	3.27248e-010	
24	4.02456e-014	1.05402e-013	3.73177e-010	
25	4.67126e-014	1.19155e-013	4.38263e-010	
26	2.56739e-014	6.66689e-014	2.39311e-010	
27	2.18159e-014	5.25691e-014	2.14072e-010	
28	4.20775e-014	1.01974e-013	4.03702e-010	
29	2.77556e-014	6.18949e-014	2.66596e-010	
30	1.19904e-014	2.67009e-014	1.20508e-010	
31	3.92186e-014	9.72555e-014	3.74598e-010	
32	5.09176e-014	1.29396e-013	4.87944e-010	
33	3.64014e-014	8.84848e-014	3.47654e-010	
34	3.68316e-014	8.95395e-014	3.51747e-010	
35	3.73868e-014	9.26481e-014	3.62661e-010	
36	3.55271e-014	8.39329e-014	3.40151e-010	. 112

# MX =