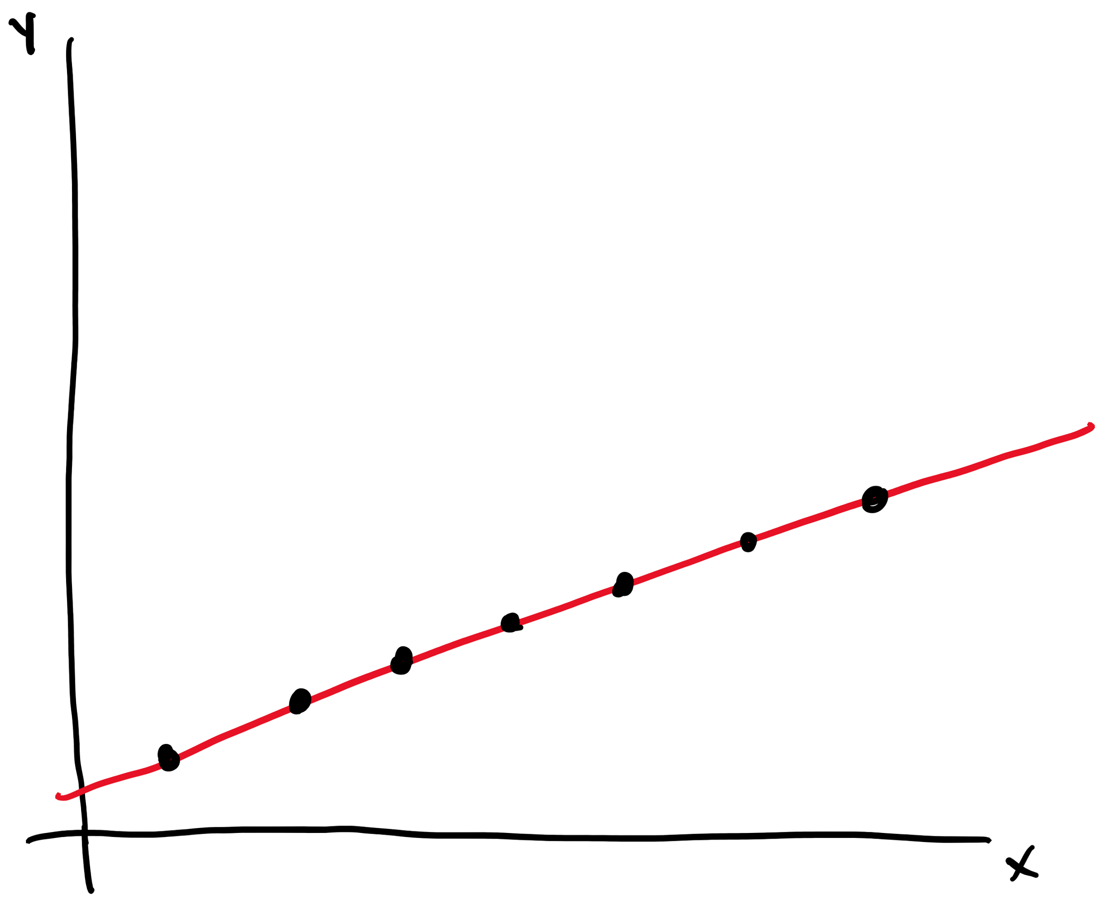
Lecture 12

Monday, June 17, 2019 2:35 PM

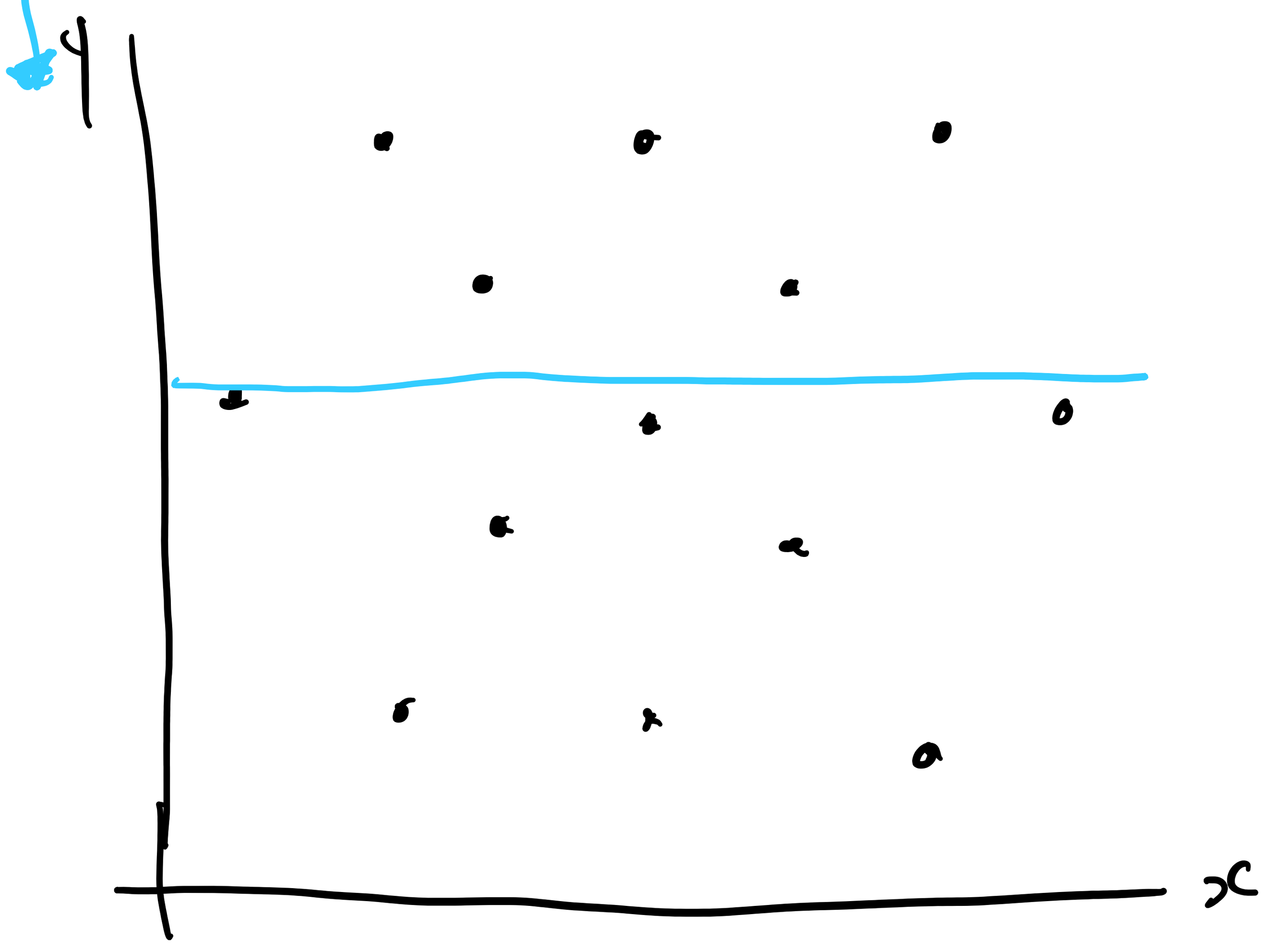


First Extreme

If the relationship between y and the x's is perfectly linear, then the fitted values will equal the observed y's exactly. When this happens our residuals are all zero and so SSE = 0. In this case the model perfectly captures the variability in the response ($R^2 = 1$).

$y_i = \beta_0 + \beta_1 x_i$

Second Extreme



When there is no relationship between the response and explanatory variables then the slope of the fitted values is 0 and so SSR = 0. In this case, the model captures none of the variability in the response ($R^2 = 0$)

$y_i = \beta_0 + \varepsilon_i, \ \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

$y_i \sim N(\beta_0, \sigma^2)$

MBA Example:

Source	df	SS	MS	F
Regression	1	0.08585	0.08585	0.3026
Error	10	2.83698	0.28370	
Total	11	2.92283		

$MSE = \hat{\sigma}^2$

$$R^2 = \frac{SSR}{SST} = \frac{0.08585}{2.92283} = 0.0294$$

∴ GMAT score explains just 2.94% of the variability in GPA scores.

Sales Example

Source	df	SS	MS	F
Regression	4	89285	22321.25	851.96
Error	10	262	26.2	
Total	14	89547		

$\hat{\sigma}^2$

$$R^2 = \frac{SSR}{SST} = \frac{89285}{89547} = 0.9971$$

∴ the model (i.e., the four explanatory variables) explain 99.71% of the variability in the response variable.

— MIDTERM COVERAGE STOPS HERE —