LINEAR REGRESSION III

11.9.2020

RECAP

* receptive fields

  * what kind of stuff in the world does a neuron (or voxel, or whatever) respond to?

  * we can estimate receptive fields by measuring responses in different conditions, and then fitting a model

RECAP

* how do we fit a stimulus–>response model?

* regression!

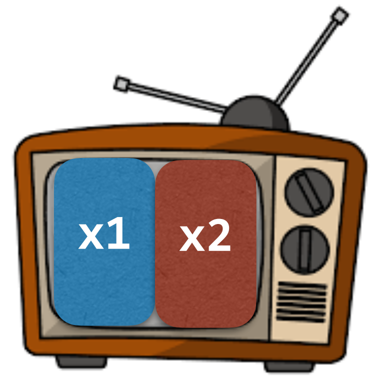
  * find weights that minimize **squared error**

  * one method uses **gradient descent**

2D EXAMPLE

subject

stimulus



* y = output of a neuron that you are measuring

* x1 = how many times **left** side of screen flashes per second

* x2 = how many times **right** side of screen flashes per second

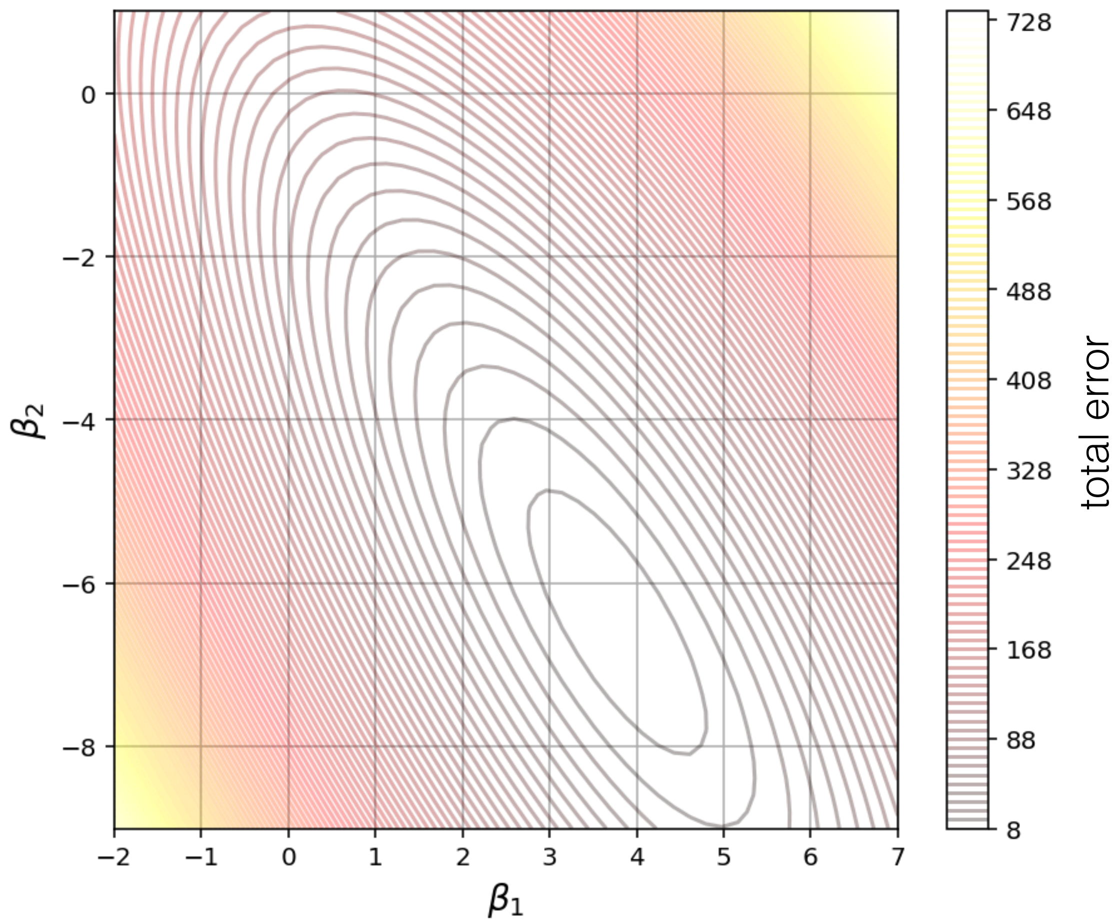
SQUARED ERROR

* **squared error** is the sum of squared differences between actual and predicted data

$$Err(\beta) = \sum_{t=1}^{T} (y_t - x_t\beta)^2$$

2D EXAMPLE

* here the error function takes two variables, *E(b1, b2)*



GRADIENT DESCENT

Squared Error: $Err(\beta) = (Y - X\beta)^2$

Gradient: $\dfrac{\partial Err(\beta)}{\partial \beta} = -X^\top (Y - X\beta)$

ANALYTIC REGRESSION

* but there's another way to find the optimal b1 and b2

* what's the shape of the error function?

  * a parabola!

ANALYTIC REGRESSION

* there is an analytic solution to the minimum of a parabola!

$$\frac{\partial Err(\beta)}{\partial \beta} = -X^\top (Y - X\beta) = 0$$

$$-X^\top Y + X^\top X\beta = 0$$

$$X^\top X\beta = X^\top Y$$

$$\beta = (X^\top X)^{-1} X^\top Y$$

ANALYTIC REGRESSION

* **np.linalg.lstsq** solves least squares regression

* it returns 4 things:

  * the regression weights (beta)

  * the residuals (final squared error)

  * the rank (we’ll talk about this later)

  * the singular values (ditto)

EVALUATING REGRESSION MODELS

* how do you know if a regression model is *good*?

* one common metric is R^2, also called the **coefficient of determination** or **variance explained**

EVALUATING REGRESSION MODELS

* R^2 = 1 - (RSS / TSS)

  * where RSS is the "residual sum of squares" (this is just another name for **squared error**)

  * and TSS is the "total sum of squares" (squared error with beta=0)

EVALUATING REGRESSION MODELS

* **R^2** can also be defined in terms of variance
    * R^2 = 1 - (var(y-y_hat) / var(y))
* what's the difference between squared error and variance?

EVALUATING REGRESSION MODELS

* suppose that we are given a matrix of variables (aka regressors) **X**, and a vector of outputs **Y**

* we fit a linear model **Y_hat = X . beta**

* then we evaluate it by computing $\mathbf{R^2}$ using **X** and **Y**

* what are the possible values of $\mathbf{R^2}$?

IN-SET VS. OUT-OF-SET EVALUATION

* evaluating a regression model using the same data that we used to train/estimate/ fit it is called *in-set evaluation*

* in-set evaluation is biased *upward*, and the amount of bias depends on the number of regressors in the model

IN-SET VS. OUT-OF-SET EVALUATION

* for example: suppose we have *N* data points and *N* regressors that are pure noise—they have no relationship to the output whatsoever

* in-set variance explained is ***EXACTLY 1.0***

* *THE MODEL IS PERFECT*

* **  **

IN-SET VS. OUT-OF-SET EVALUATION

* instead, what if you split up your X and Y into “training” and “test” sets?

* you could fit your regression model using (X_trn, Y_trn), and then test how well it works on (X_test, Y_test)!

* is $\mathbf{R}^2$ biased in this case? What possible values can it take?

END