

« Previous (/course/cs357-f15/flow-session/74262/3/)

下一页 » (/course/cs357-f15/flow-session/74262/5/)

跳转至最末 »

1 2 3 (/course/cs357-f15/flow-session/74262/0/) (/course/cs357-f15/flow-session/74262/1/)
4 (/course/cs357-f15/flow-session/74262/2/) (/course/cs357-f15/flow-
5 6 session/74262/3/) (/course/cs357-f15/flow-session/74262/5/)

Solving for weights

We have made a hypothesis, that there is a linear relationship in the parameters. Let's do two things: 1. Solve for the linear relationship 2. Validate the data with another data set

So first:

1. Solve for the linear representation weights using QR factorizations of the A matrices built from the *breast-cancer-train.dat* data set. You should use your b built from the *breast-cancer-train.dat* data set as well.
 - Call the weights `weights_linear`.
2. See how well your weights predict whether a tumor is malignant or benign by simply performing matrix-vector multiplication of the linear and quadratic representation of the *breast-cancer-validate.dat* data set with `weights_linear`. *Why?*

This will produce a prediction vector p . For each entry i of p , If $p[i] > 0$, we will assume the i th person has a malignant tumor. If $p[i] \leq 0$, then the i th person has a benign tumor.

- Compare your p for the linear and quadratic representation to the actual b (Use the $p[i] > 0$, $p[i] \leq 0$ rule given above).
- Store the number of false-positives and false-negatives for both representations. A false-positive is when your model indicated that a tumor was malignant, when it was not. A false-negative is when your model fails to correctly identify a tumor as malignant.
- Use a bar-graph to plot the number of false-positives fp and false-negatives fn for each representation.