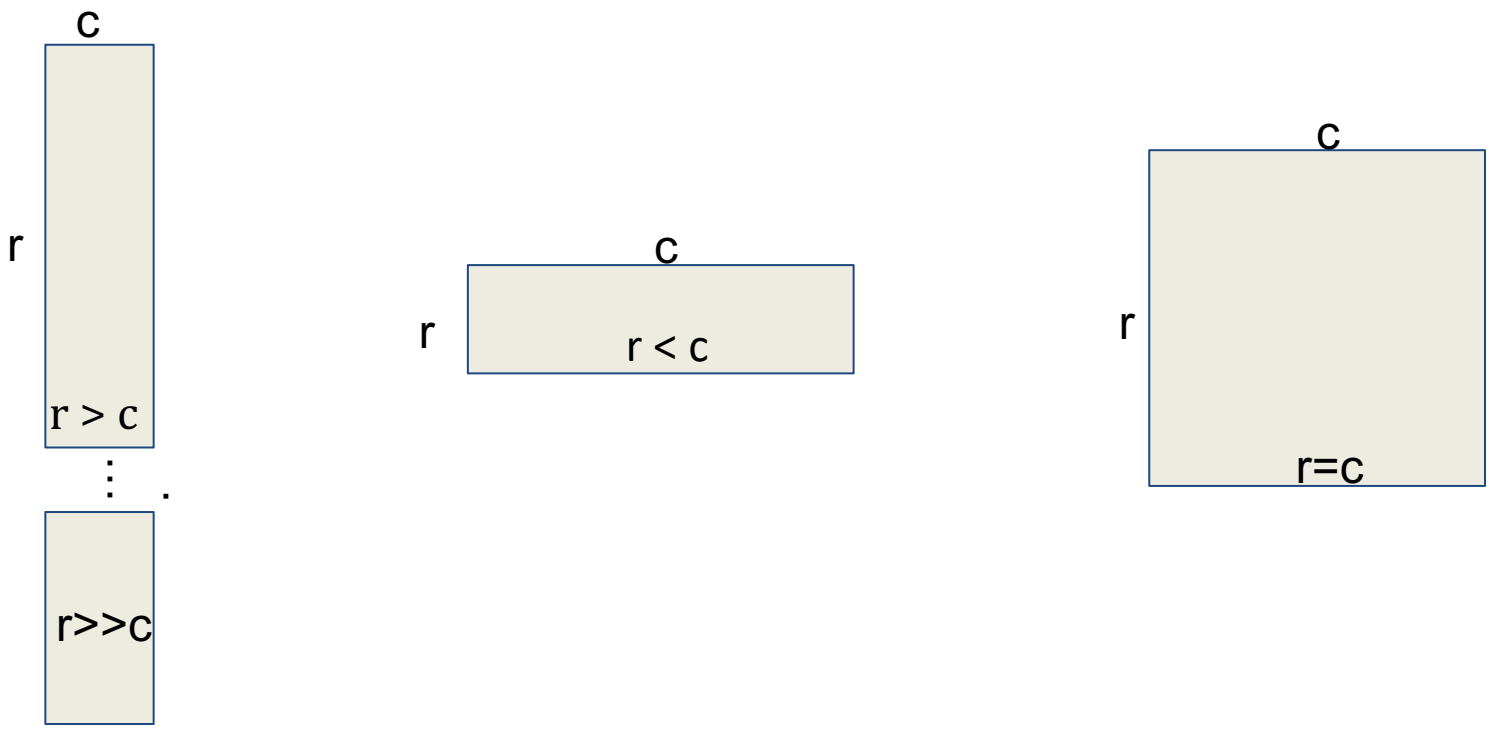
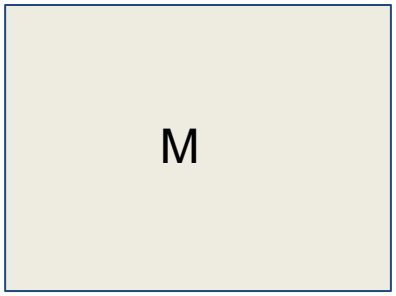
Introduction to Three* Matrix Structures

*Interpreted 4 ways

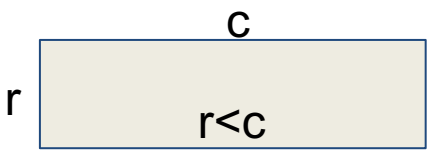


Framing



M or X or df or data or observations

L or Y or target



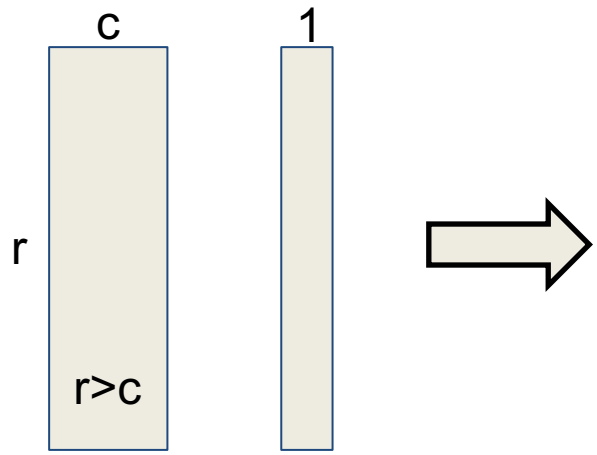
[m,n] or [r,c] or [x,y] or a_ij

```
M = [[1,2,3,4],
     [4,5,6,7],
     [7,8,9,10]]
```

```
L = [1,2,3,4]
```

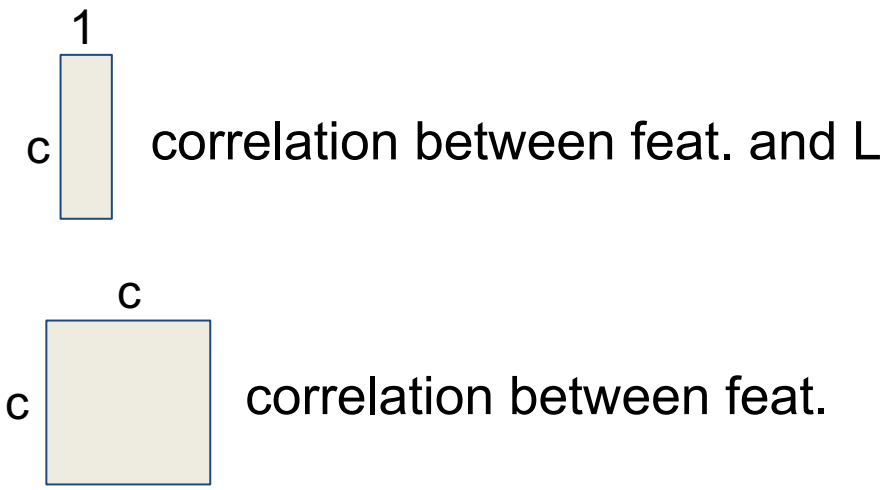
Tall and Skinny Matrices

Explicitly calculate correlation between columns



M
X
df
data
observations

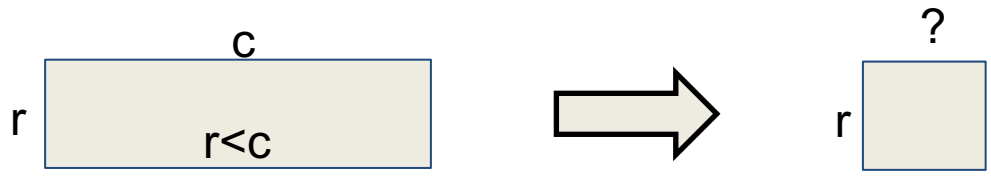
L
Y
target



Classic Question: Do we have enough data to estimate correlation between features and labels?

Short and Fat Matrices

Approximate the correlation between columns



Simplifying assumption: a good model doesn't require all provided features

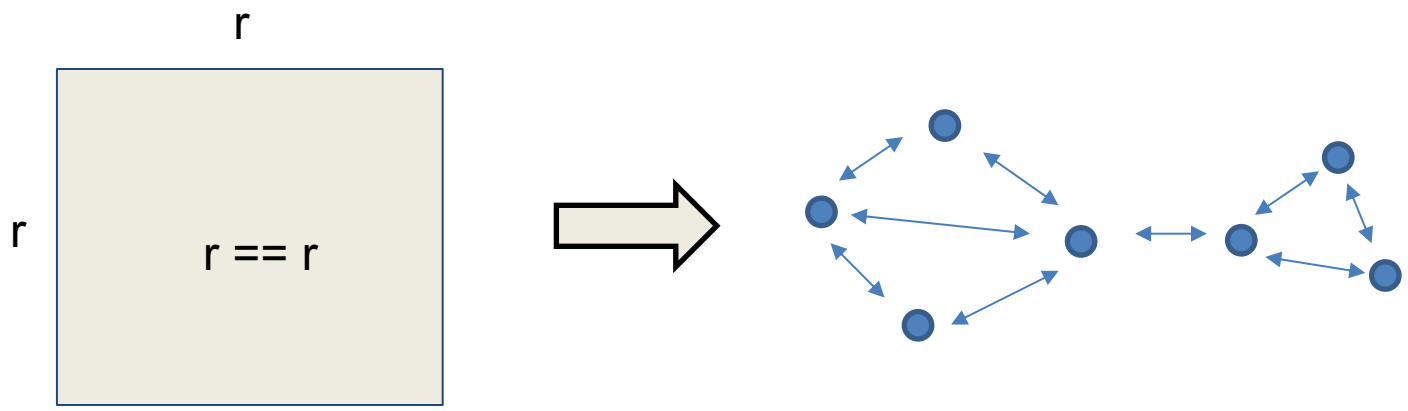
PCA: preforms lossless compression on M independent of L

Lasso, Ridge Regression: utilizes L to identify features that are significant

Random Forest: utilizes L to identify features that are significant

Square

Network analysis, Graph Representations



Every a_ij entry represents the relationship between item i and item j.

i.e. clustering, node centrality, propagation models

Really Tall and Skinny Matrices

Explicitly calculate higher order correlations between columns