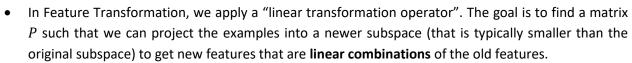
Scan me

# UL04. Feature Transformation NEEDS REWORK

#### What is Feature Transformation?

- The problem of pre-processing a set of features to create a new (more compact) feature set, while retaining as much (relevant/useful) information as possible.
- Feature Selection is a subset of Feature Transformation, where the preprocessing is literally extracting a subset of the features.



- Why?
  - We combine features together hoping to eliminate false positives/negatives.

## Principal Components Analysis:

- Principal Components Analysis will transform the features set by:
  - Finding the direction (vector) that maximizes variance. The is called the Principal Component.
  - Finding directions that are orthogonal to the Principal Component.
- PCA is a global algorithm
- PCA gives the ability to do reconstruction, because we don't lose information.

### **Independent Components Analysis:**

- ICA attempts to maximize independence. It tries to find a linear transformation of the feature space, such that each of the individual new features are mutually statistically independent.
  - The mutual information between any two random features equals zero:

$$I(y_i; y_i) = 0$$

- The mutual information between the new features set and the old features set is as high as possible:

$$I(Y;X) = \uparrow \uparrow$$

## Random Components Analysis:

- Similar to Principal Components Analysis, but instead of generating directions that maximize variance, it generates random directions.
- It captures some of the correlations that works well with classification settings.
- It's faster than PCA and ICA.

# Linear Discriminant Analysis:

• Linear Discriminant Analysis finds a projection that discriminates based on the label.