DATA SCIENCE
PCA

AGENDA

- Last Class
- Linear Algebra Review
- Dimension Reduction
- PCA
- Kernel PCA
- Python Practice

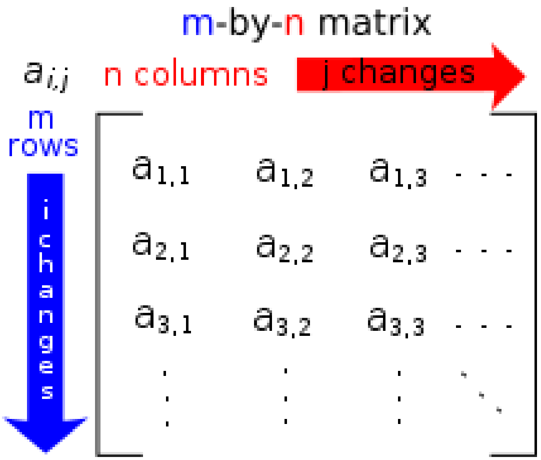
LINEAR ALGEBRA

Q: What is a matrix?

Q: What is a matrix?

A: Rectangular array of rows and columns.

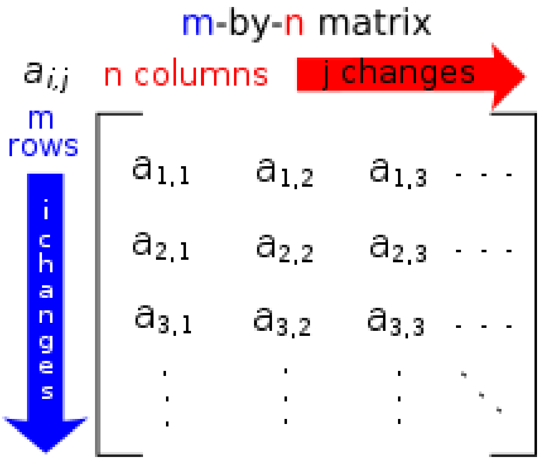
An m x n matrix has m rows and n columns



LINEAR ALGEBRA

Q: What is a matrix?

A: Rectangular array of rows and columns.
An m x n matrix has m rows and n columns



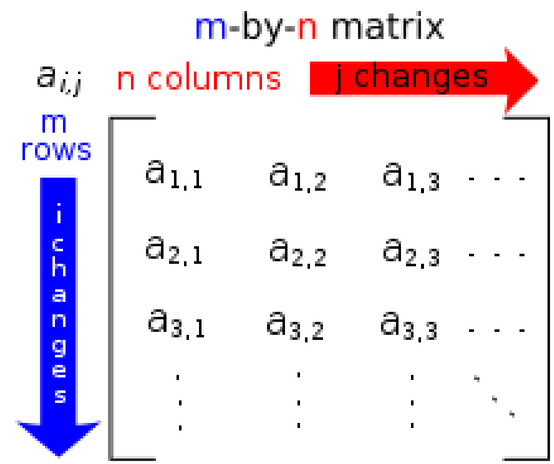
Q: How do you do matrix vector multiplication?

LINEAR ALGEBRA

Q: What is a matrix?

A: Rectangular array of rows and columns.

An m x n matrix has m rows and n columns



Q: How do you do matrix vector multiplication?

A:

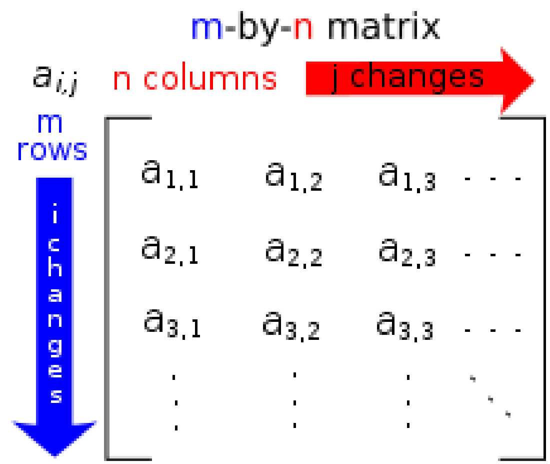
$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{bmatrix}$$

LINEAR ALGEBRA

Q: What is a matrix?

A: Rectangular array of rows and columns.

An m x n matrix has m rows and n columns



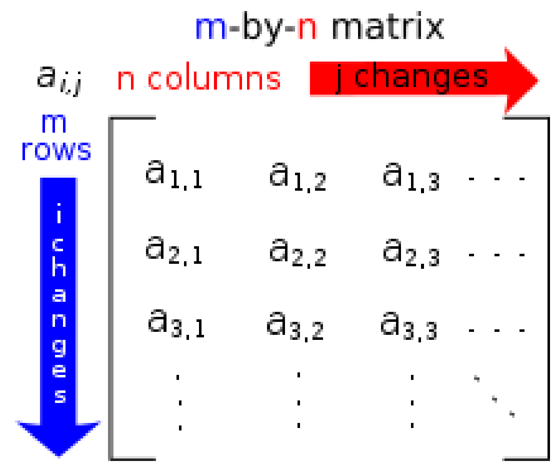
Q: How do you do matrix vector multiplication?

A:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{bmatrix}$$

Q: How do you do matrix, matrix multiplication?

LINEAR ALGEBRA

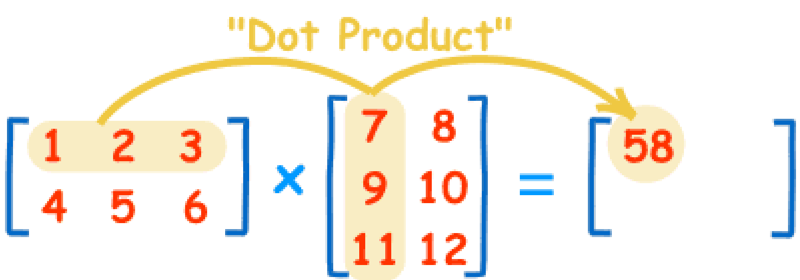


Q: What is a matrix?
A: Rectangular array of rows and columns.
An m x n matrix has m rows and n columns

Q: How do you do matrix vector multiplication?
A:

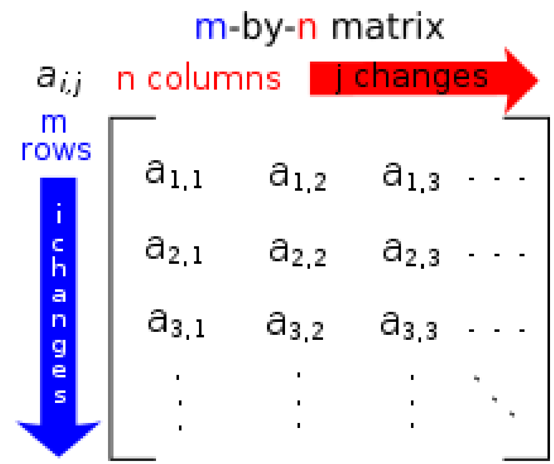
$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{bmatrix}$$

Q: How do you do matrix, matrix multiplication?
A:



LINEAR ALGEBRA

Q: What is a matrix?
A: Rectangular array of rows and columns.
An m x n matrix has m rows and n columns

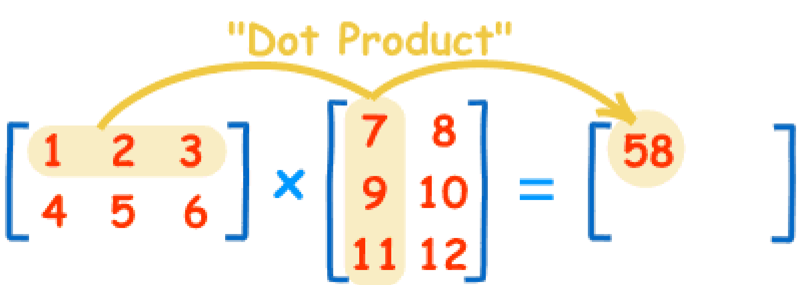


Q: How do you do matrix vector multiplication?

A: $$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{bmatrix}$$

Q: How do you do matrix, matrix multiplication?

A:



- A is an mxn matrix
- B is an nxp matrix
- AxB is an mxp matrix
- Note:# columns of A = # rows of B

DIMENSIONALITY REDUCTION

Dimensionality reduction is unsupervised learning

Q: What does this mean?

DIMENSIONALITY REDUCTION

Dimensionality reduction is unsupervised learning

Q: What does this mean?
A: We don’t have a y value. We’re trying to find structure, relationships, etc

DIMENSIONALITY REDUCTION

Q: What is dimensionality reduction?

DIMENSIONALITY REDUCTION

Q: What is dimensionality reduction?

A: A set of techniques used to reduce the number of dimensions in our data set

DIMENSIONALITY REDUCTION

Q: What is dimensionality reduction?
A: A set of techniques used to reduce the number of dimensions in our data set

Q: Haven’t we done this before with feature selection?

DIMENSIONALITY REDUCTION

Q: What is dimensionality reduction?
A: A set of techniques used to reduce the number of dimensions in our data set

Q: Haven't we done this before with feature selection?
- Yes we reduced the dimensions by selecting a subset of the features
- In this context, we're going to transform the data into a lower dimensional space

DIMENSIONALITY REDUCTION

Q: What is dimensionality reduction?
A: A set of techniques used to reduce the number of dimensions in our data set

Q: Haven't we done this before with feature selection?
- Yes we reduced the dimensions by selecting a subset of the features
- In this context, we're going to transform the data into a lower dimensional space

Q: What are some advantages?

DIMENSIONALITY REDUCTION

Q: What is dimensionality reduction?
A: A set of techniques used to reduce the number of dimensions in our data set

Q: Haven't we done this before with feature selection?

- Yes we reduced the dimensions by selecting a subset of the features
- In this context, we're going to transform the data into a lower dimensional space

Q: What are some advantages?

- Reduced computation and storage
- Visualization
- Reduce multi-collinearity
- Avoid overfitting (Reduce noise while keeping information)

PRINCIPAL COMPONENT ANALYSIS

Q: What is principal components analysis?

PRINCIPAL COMPONENT ANALYSIS

Q: What is principal components analysis?

A: A linear dimensionality reduction method

PRINCIPAL COMPONENT ANALYSIS

Q: What is principal components analysis?
A: A linear dimensionality reduction method

Q: What does PCA do?

PRINCIPAL COMPONENT ANALYSIS

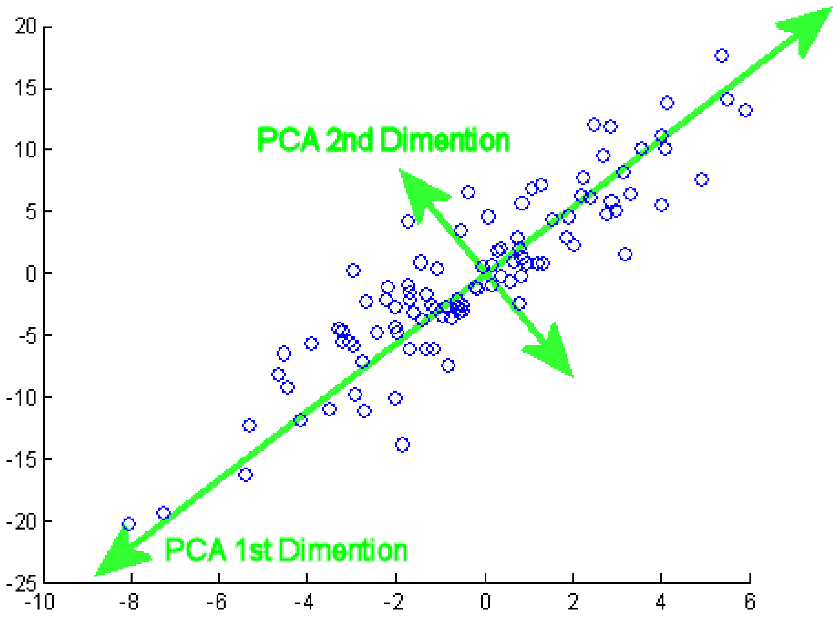
Q: What is principal components analysis?
A: A linear dimensionality reduction method

Q: What does PCA do?
A: Find the directions of maximum variance in the data

PRINCIPAL COMPONENT ANALYSIS

Q: What is principal components analysis?
A: A linear dimensionality reduction method

Q: What does PCA do?
A: Find the directions of maximum variance in the data



PRINCIPAL COMPONENT ANALYSIS

Let’s dig a little deeper

Q: What is variance?

PRINCIPAL COMPONENT ANALYSIS

Let's dig a little deeper

Q: What is variance?

A: A measure of how spread out the data is $\mathrm{Var}(X) = \mathrm{E}\left[(X-\mu)^2\right]$.

PRINCIPAL COMPONENT ANALYSIS

Let's dig a little deeper

Q: What is variance?

A: A measure of how spread out the data is $\mathrm{Var}(X) = \mathrm{E}\left[(X - \mu)^2\right].$

Q: What is covariance?

PRINCIPAL COMPONENT ANALYSIS

Let's dig a little deeper

Q: What is variance?

A: A measure of how spread out the data is $\mathrm{Var}(X) = \mathrm{E}\left[(X-\mu)^2\right].$

Q: What is covariance?

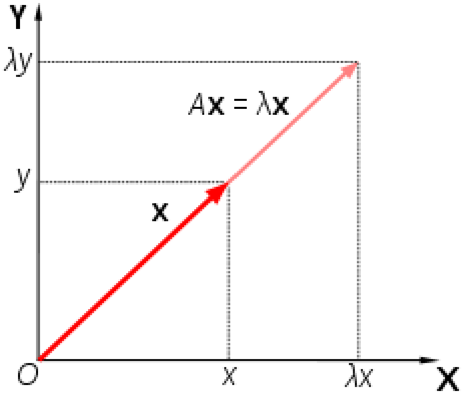
A: Variance only measures this on one dimension. Covariance is a measure of how much different dimensions change together

$$\sigma(X,Y) = \mathrm{E}\left[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])\right],$$

PRINCIPAL COMPONENT ANALYSIS

Anyone know about Eigenvalues and Eigenvectors?

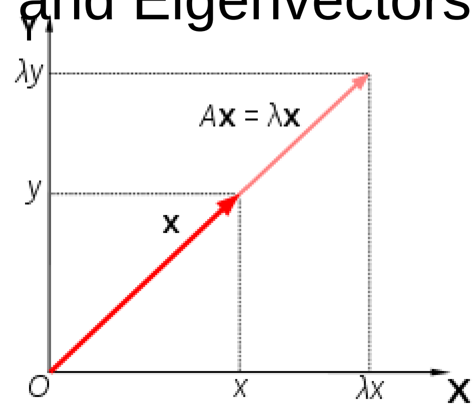
$$A\mathbf{v} = \lambda\mathbf{v},$$



PRINCIPAL COMPONENT ANALYSIS

Anyone know about Eigenvalues and Eigenvectors?

$$A\mathbf{v} = \lambda\mathbf{v},$$



It turns out that the eigenvectors of the covariance matrix are the principal components where the eigenvalues tell us the magnitude

If A is a square nxn matrix with n linearly idependent eigenvectors then we can factorize A as such

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$$

Where Q is made up of the eigenvectors and LAMBDA a diagonal matrix of eigenvalues

PRINCIPAL COMPONENT ANALYSIS

Let’s dig a little deeper

Q: What is variance?

A: A measure of how spread out the data is $\operatorname{Var}(X) = \mathrm{E}\left[(X-\mu)^2\right].$

Q: What is covariance?

A: Variance only measures this on one dimension. Covariance is a measure of how much different dimensions change together

$$\sigma(X,Y) = \mathrm{E}\left[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])\right],$$

Covariance matrix: $$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \\ \sigma_{n1} & \sigma_{n2} & & \sigma_{nn} \end{pmatrix}$$

PRINCIPAL COMPONENT ANALYSIS

Steps:

- Standardize the dataset
- Construct the covariance matrix
- Find eigenvalues and eigenvectors
- Select k eigenvectors corresponding to the k largest eigenvalues
- Construct a projection matrix W from the top k eigenvectors
- Now can transform the d-dimensional data X into k-dimensions using W

PRINCIPAL COMPONENT ANALYSIS

Steps:

- Standardize the dataset
- Construct the covariance matrix
- Find eigenvalues and eigenvectors
- Select k eigenvectors corresponding to the k largest eigenvalues
- Construct a projection matrix W from the top k eigenvectors
- Now can transform the d-dimensional data X into k-dimensions using W

Why do we need to standardize?

PRINCIPAL COMPONENT ANALYSIS

Steps:

- Standardize the dataset
- Construct the covariance matrix
- Find eigenvalues and eigenvectors
- Select k eigenvectors corresponding to the k largest eigenvalues
- Construct a projection matrix W from the top k eigenvectors
- Now can transform the d-dimensional data X into k-dimensions using W

Why do we need to standardize?
What the variance of range(10)?

PRINCIPAL COMPONENT ANALYSIS

Steps:

- Standardize the dataset
- Construct the covariance matrix
- Find eigenvalues and eigenvectors
- Select k eigenvectors corresponding to the k largest eigenvalues
- Construct a projection matrix W from the top k eigenvectors
- Now can transform the d-dimensional data X into k-dimensions using W

Why do we need to standardize?
What the variance of range(10)? 8.25
What is the variance of 10*range(10)?

PRINCIPAL COMPONENT ANALYSIS

Steps:

- Standardize the dataset
- Construct the covariance matrix
- Find eigenvalues and eigenvectors
- Select k eigenvectors corresponding to the k largest eigenvalues
- Construct a projection matrix W from the top k eigenvectors
- Now can transform the d-dimensional data X into k-dimensions using W

Why do we need to standardize?
What the variance of range(10)? 8.25
What is the variance of 10*range(10)? 825

KERNEL PRINCIPAL COMPONENT ANALYSIS

PCA is a linear projection of the data.

Q: How can we make this nonlinear?

KERNEL PRINCIPAL COMPONENT ANALYSIS

PCA is a linear projection of the data.

Q: How can we make this nonlinear?

A: Similar to SVMs we use a kernel

Questions?