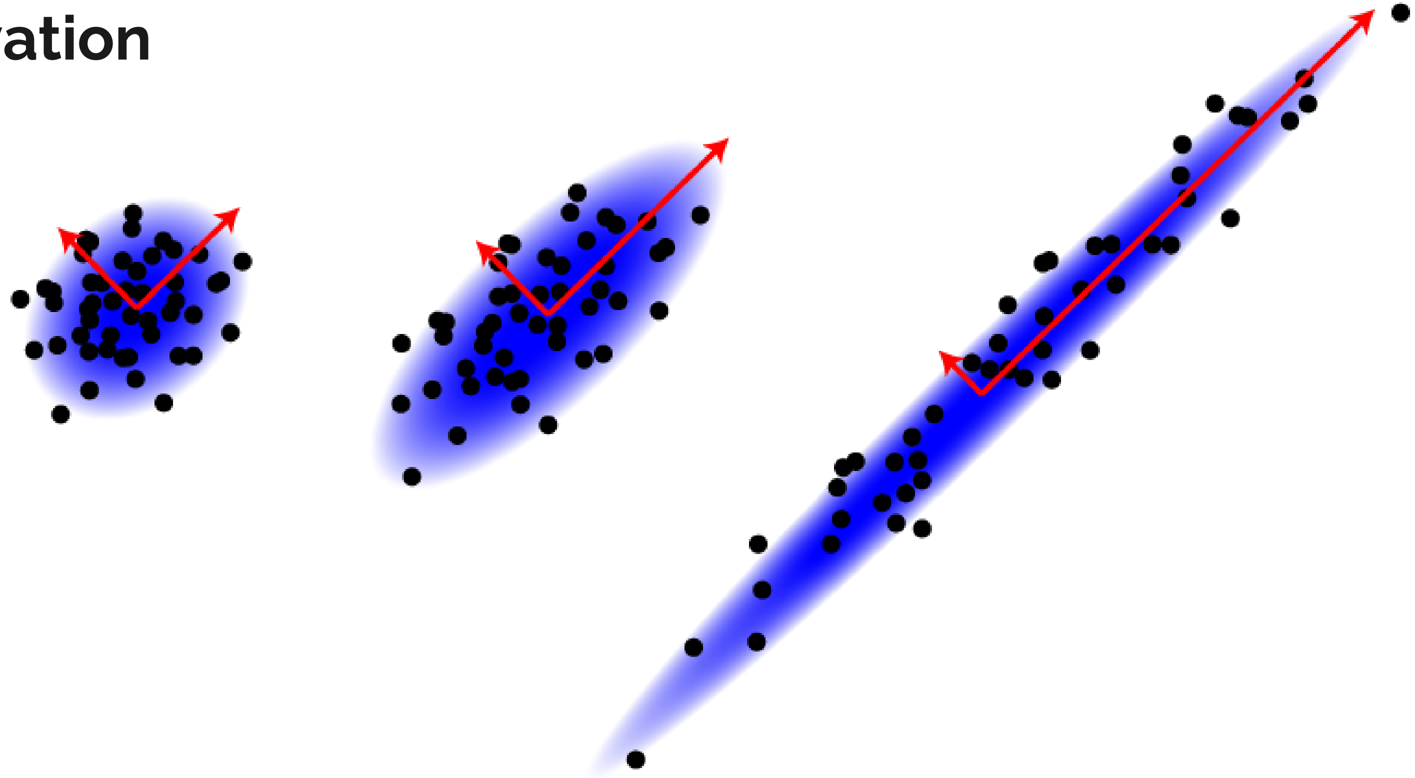
An Introduction to Principal Component Analysis

By William Gebhardt

Neural Adaptive Computing Laboratory

Motivation

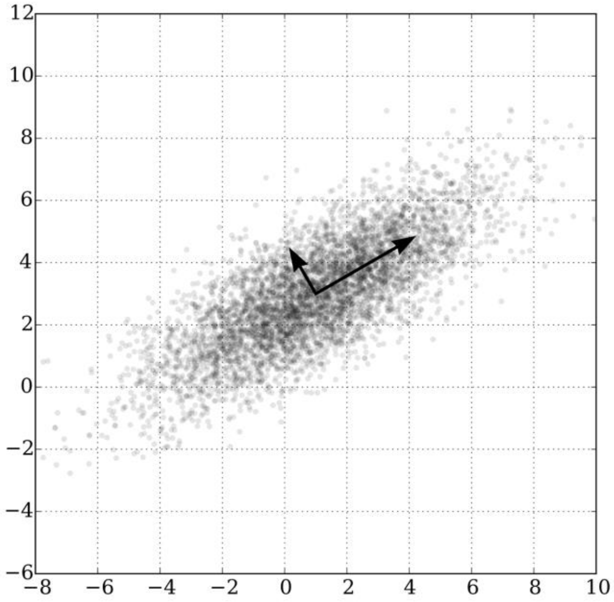


History

- Invented in 1901 by Karl Pearson
- In the 1930s Harold Hotelling independently developed a similar approach
- There are many different names for practically the same method
  - Karhunen Loève Transform (Signal Processing)
  - Hotelling Transform (Multivariate Quality Control)
  - Proper Orthogonal Decomposition (Mechanical Engineering)
  - Singular Value Decomposition (Linear Algebra)
  - Many More

What is PCA?

- High dimension/feature reduction
    - Breaks down a collection of data points into principal components
- Purely linear transformations
- Only a small portion of the components are needed to cover much of the variance



Background Math

Zero-mean and Standardization

Zero-mean

- Shifts the mean to have the values centered at zero

$$X - \bar{X}$$

Standardization

- Shifts all the points to have unit variance

$$\frac{X}{\sigma_X}$$

Z-score equation → $$z_X = \frac{X - \bar{X}}{\sigma_X}$$

Independent Vectors and Matrix Inverses

Independent Vectors

- Two vectors that can not be made equal by multiplying by a scalar

$$\{1, 1, 2\}$$
$$\{2, -1, 1\}$$

Matrix Inverse

- Equivalent to the reciprocal of a matrix

$$AA^{-1} = I$$
$$A^{-1} = \frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

The Determinant

- Only exists for a square matrix
- Produces a scalar value from a matrix
- Determinant of a product of matrices is the product of their determinants

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb$$

$$|A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & i \end{vmatrix}$$

Covariance

- The measure of the joint variability of two random variables
- The Covariance Matrix of vectors is a symmetric matrix of the covariance between dimensions of vectors
    - The row and column (i, j) correspond to the i'th and j'th dimension

$$\mathrm{cov}(\mathrm{X},\, \mathrm{Y}) = \mathrm{E}[(X - \mathrm{E}(X))(Y - \mathrm{E}(Y))]$$

Eigenvalues and Eigenvectors

Base Equations

- X: A nonzero vector existing in n-dimensional space known as an eigenvector
- A: A linear transformation applied to X represented as a square matrix
- λ: A scalar value known as an eigenvalue

$$X \in R^n \neq 0$$

$$AX = \lambda X$$

Finding Eigenvalues

- Rework our base equation into a homogeneous system
- Cramer's Rule states that a linear system of equations has nontrivial solutions iff the determinant vanishes
- This is known as the characteristic equation of A

$$AX = \lambda X$$

$$AX - \lambda X = 0$$

$$(A - \lambda I)X = 0$$

$$\det(A - \lambda I) = 0$$

Example

$$A = \begin{bmatrix} -2 & -2 & 4 \\ -4 & 1 & 2 \\ 2 & 2 & 5 \end{bmatrix}$$

$$det(A - \lambda I) = 0$$

$$det \begin{bmatrix} -2 & -2 & 4 \\ -4 & 1 & 2 \\ 2 & 2 & 5 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = 0$$

$$det \begin{bmatrix} -2-\lambda & -2 & 4 \\ -4 & 1-\lambda & 2 \\ 2 & 2 & 5-\lambda \end{bmatrix} = 0$$

$$(-2-\lambda)[(1-\lambda)(5-\lambda)-(2)(2)]-(-4)[(-2)(5-\lambda)-(4)(2)]+(2)[(-2)(2)-(4)(1-\lambda)]=0$$

$$-\lambda^3+4\lambda^2+27\lambda-90=0$$
$$\lambda^3-4\lambda^2-27\lambda+90=0$$
$$(\lambda-3)(\lambda^2-\lambda-30)=0$$
$$(\lambda-3)(\lambda+5)(\lambda-6)=0$$

$$\lambda = 3$$

$$\begin{bmatrix} -2-\lambda & -2 & 4 \\ -4 & 1-\lambda & 2 \\ 2 & 2 & 5-\lambda \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = 0$$

$$\begin{bmatrix} -5 & -2 & 4 \\ -4 & -2 & 2 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = 0$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$$

Recap

- If we solve for the nontrivial values of λ we will get up to *n* different eigenvalues (3, -5, 6: in the example)
- When we plug these values into the base equation we will produce n systems of equations to solve
- The results are the eigenvectors corresponding to the eigenvalues