GTDM – 2019/20

- Determinants
- Eigenvalues/Eigenvectors

Based on Linear Algebra and Its Applications, David C. Lay, Steven R. Lay, and Judi J. McDonald, PEARSON 5^th ed.

The Determinant of a Matrix

With each $n \times n$ matrix A it is possible to associate a scalar, det(A), whose value will tell us whether the matrix is nonsingular (i.e. invertible).
Geometrically, this scalar is related to the volume of the "rectangular region" with the rows of the matrix as sides.

Case 1: 1 × 1 Matrices If $A = (a)$ is a 1×1 matrix, then A will have a multiplicative inverse if and only if $a \neq 0$. Thus, if we define $\det(A) = a$, then A will be nonsingular if and only if $\det(A) \neq 0$.

Case 2: 2 × 2 Matrices Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

We will check if A is nonsingular using equivalent form. Then, if $a_{11} = 0$, we perform the following operations:

1. Multiply the second row of A by a_{11}

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{11}a_{21} & a_{11}a_{22} \end{pmatrix}$$

2. Subtract a_{21} times the first row from the new second row

$$\begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{11}a_{22} - a_{21}a_{12} \end{pmatrix}$$

Since $a_{11} \neq 0$, the resulting matrix has rank equal to 2 if and only if

$$a_{11}a_{22} - a_{21}a_{12} \neq 0$$

If $a_{11} = 0$, we can switch the two rows of A. The resulting matrix $A = \begin{bmatrix} a_{21} & a_{22} \\ 0 & a_{12} \end{bmatrix}$

Is non singular if and only if $a_{21}a_{12} \neq 0$. This requirement is equivalent to the previous condition (when $a_{11} \neq 0$).
Thus, if A is any 2 × 2 matrix and we define

$$\det(A) = a_{11}a_{22} - a_{12}a_{21}$$

then A is nonsingular if and only if $\det(A) \neq 0$.

Remark The value $|a_{11}a_{22} - a_{21}a_{12}|$ is equal to the area of the parallelogram determined by the vectors (a_{11}, a_{12}) and (a_{21}, a_{22}).

Notation

We can refer to the determinant of a specific matrix by enclosing the array between vertical lines. For example, if

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Then

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$$

represents the determinant of A.

Case 3: 3 × 3 Matrices We can test whether a 3 × 3 matrix is nonsingular by performing row operations. to see if To carry out the elimination in the first column of an arbitrary 3 × 3 matrix *A*, let us first assume that $a_{11} \neq 0$. The elimination can then be performed by subtracting a_{21}/a_{11} times the first row from the second and a_{31}/a_{11} times the first row from the third:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} & \dfrac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}} \\ 0 & \dfrac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}} & \dfrac{a_{11}a_{33} - a_{31}a_{13}}{a_{11}} \end{pmatrix}$$

The matrix on the right will be nonsingular if and only if

$$a_{11} \begin{vmatrix} \dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} & \dfrac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}} \\ \dfrac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}} & \dfrac{a_{11}a_{33} - a_{31}a_{13}}{a_{11}} \end{vmatrix} \neq 0$$

this condition can be simplified to (the algebra is somewhat messy….),

$$det(A) = a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22} \neq 0$$

What if $a_{11} = 0$? Consider the following possibilities:

(i) $a_{11} = 0$, $a_{21} \neq 0$

(ii) $a_{11} = a_{21} = 0$, $a_{31} \neq 0$

(iii) $a_{11} = a_{21} = a_{31} = 0$

which however lead to a similar condition

We would now like to define the determinant of an $n \times n$ matrix. To see how to do this, note that the determinant of a 2×2 matrix, can be defined in terms of the two 1×1 matrices:

$$M_{11} = (a_{22}) \text{ and } M_{12} = (a_{21})$$

The matrix M_{11} is formed from A by deleting its first row and first column, and M_{12} is formed from A by deleting its first row and second column. The determinant of A can be expressed in the form

$$\det(A) = a_{11}a_{22} - a_{12}a_{21} = a_{11}\det(M_{11}) - a_{12}\det(M_{12})$$

For a 3×3 matrix A, we can rewrite det(A) in the form

$$\det(A) = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22})$$

For $j = 1, 2, 3$, let M_{1j} denote the 2×2 matrix formed from A by deleting its first row and jth column. The determinant of A can then be represented in the form

$$\det(A) = a_{11}\det(M_{11}) - a_{12}\det(M_{12}) + a_{13}\det(M_{13})$$

where

$$M_{11} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}, \quad M_{12} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}, \quad M_{13} = \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Definition

Let $A = (a_{ij})$ be an $n \times n$ matrix and let M_{ij} denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting the row and column containing a_{ij}. The determinant of M_{ij} is called the **minor** of a_{ij}. We define the **cofactor** A_{ij} of a_{ij} by

$$A_{ij} = (-1)^{i+j} \det(M_{ij})$$

Example. In view of this definition, for a 2 × 2 matrix A, we have $\det(A) = a_{11}A_{11} + a_{12}A_{12}$ ($n = 2$)

Remark. Note that we could also write $\det(A) = a_{21}(-a_{12}) + a_{22}a_{11} = a_{21}A_{21} + a_{22}A_{22}$ which expresses det(A) in terms of the entries of the second row of A and their cofactors. Actually, there is no reason that we must expand along a row of the matrix; the determinant could just as well be represented by the cofactor expansion along one of the columns.

Example. For a 3 × 3 matrix A, we have $\det(A) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$

$$\begin{vmatrix} 2 & 5 & 4 \\ 3 & 1 & 2 \\ 5 & 4 & 6 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 \\ 4 & 6 \end{vmatrix} - 5\begin{vmatrix} 3 & 2 \\ 5 & 6 \end{vmatrix} + 4\begin{vmatrix} 3 & 1 \\ 5 & 4 \end{vmatrix} = 2(6-8) - 5(18-10) + 4(12-5)$$
$$= -16$$

Definition

The **determinant** of an $n \times n$ matrix A, denoted $\det(A)$, is a scalar associated with the matrix A that is defined inductively as

$$\det(A) = \begin{cases} a_{11} & \text{if } n = 1 \\ a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} & \text{if } n > 1 \end{cases}$$

where

$$A_{1j} = (-1)^{1+j}\det(M_{1j}) \qquad j = 1, \ldots, n$$

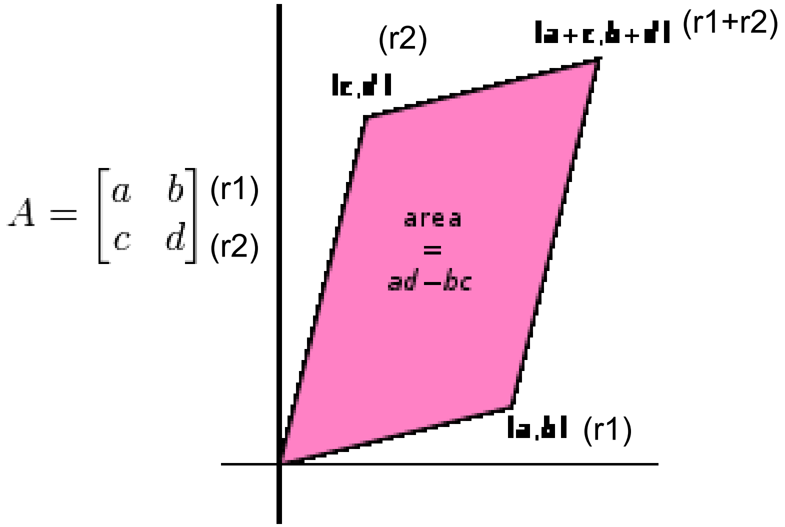
are the cofactors associated with the entries in the first row of A.

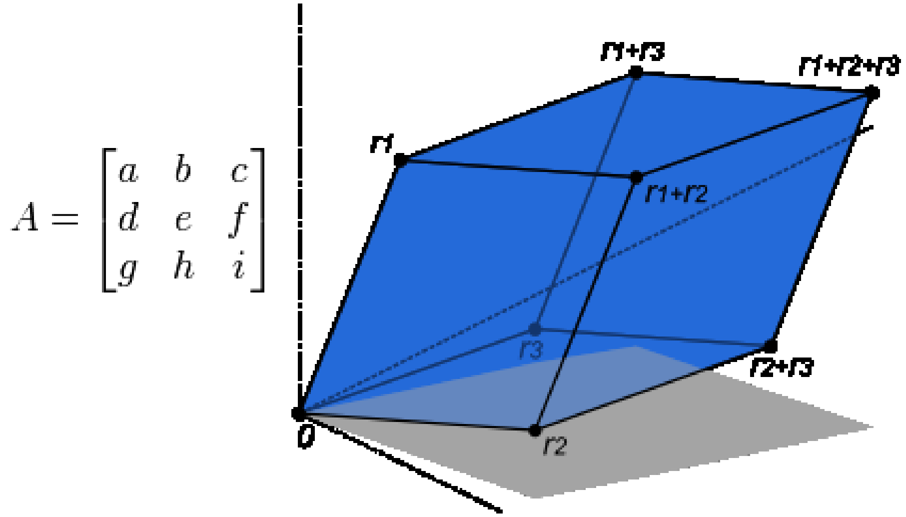
Theorem 2.1.1 *If A is an $n \times n$ matrix with $n \geq 2$, then $\det(A)$ can be expressed as a cofactor expansion using any row or column of A.*

$$\begin{aligned} \det(A) &= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} \end{aligned}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, n$.

Matrix Determinant

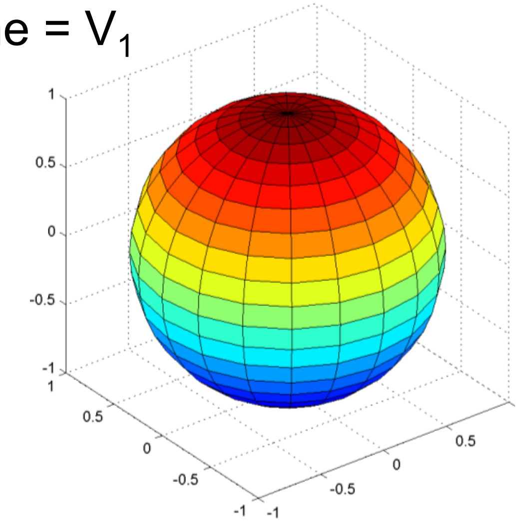




- The determinant is the “volume” of a matrix
- Actually the volume of a parallelepiped formed from its row vectors
  - Also the volume of the parallelepiped formed from its column vectors

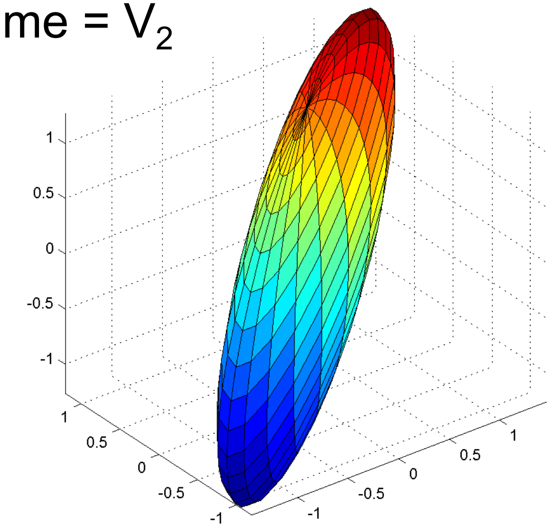
Matrix Determinant: Another Perspective

Volume = V_1



$$\begin{bmatrix} 0.8 & 0 & 0.7 \\ 1.0 & 0.8 & 0.8 \\ 0.7 & 0.9 & 0.7 \end{bmatrix}$$

Volume = V_2



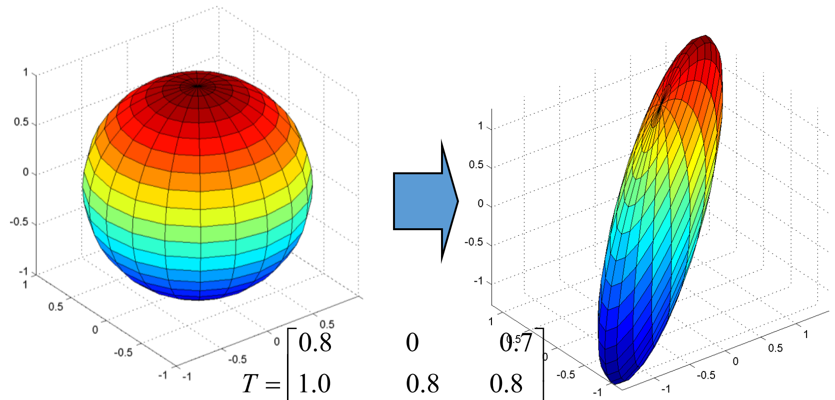
- The determinant is the ratio of N-volumes
  - If V_1 is the volume of an N-dimensional sphere “O” in N-dimensional space
    - O is the complete set of points or vertices that specify the object
  - If V_2 is the volume of the N-dimensional ellipsoid specified by A*O, where A is a matrix that transforms the space
  - $|A| = V_2 / V_1$

Matrix Inversion

- A matrix transforms an N-dimensional object to a different N-dimensional object

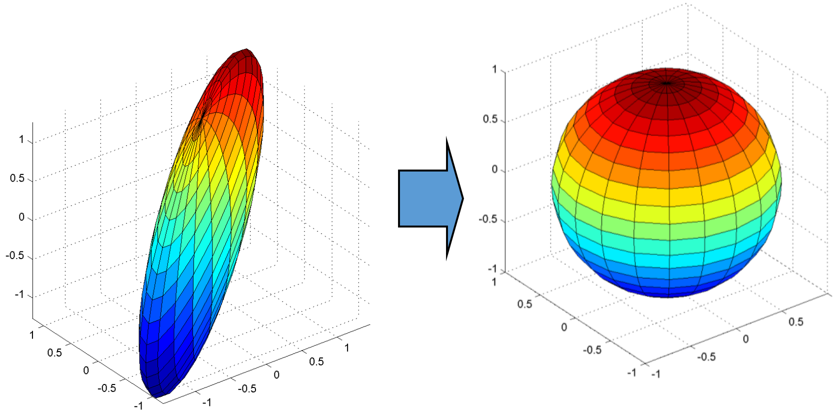
- What transforms the new object back to the original?
  - The *inverse transformation*

- The inverse transformation is called the matrix inverse



$$T = \begin{bmatrix} 0.8 & 0 & 0.7 \\ 1.0 & 0.8 & 0.8 \\ 0.7 & 0.9 & 0.7 \end{bmatrix}$$

$$Q = \begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} = T^{-1}$$



Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If $ad - bc = 0$, then A is not invertible.

The quantity ad-bc is the **determinant** of A det A

Example. $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$. $\longrightarrow$ $A^{-1} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}$

If A is an invertible $n \times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

Dynamical Systems and Spotted Owls

A first step in studying the population dynamics of the spotted owls is to model the population at yearly intervals, at times denoted by k=0, 1, 2, … Usually, one assumes that there is a 1:1 ratio of males to females in each life stage and counts only the females. The population at year k can be described by a vector

$$\mathbf{x}_k = (j_k, s_k, a_k),$$

where j_k, s_k, and a_k are the numbers of females in the juvenile, subadult, and adult stages, respectively. Using actual field data from demographic studies, R. Lamberson and co-workers considered the following *stage-matrix model*:

$$\begin{bmatrix} j_{k+1} \\ s_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \begin{bmatrix} j_k \\ s_k \\ a_k \end{bmatrix}$$

$$\boldsymbol{x}_{k+1} = A \cdot \boldsymbol{x}_k$$

$$\text{where} \begin{cases} j_k = \text{number of juvenile spotted owls at year } k \\ s_k = \text{number of subadult spotted owls at year } k \\ a_k = \text{number of adult spotted owls at year } k \end{cases}$$

so that

$$\begin{aligned} j_{k+1} &= 0.33\, a_k \\ s_{k+1} &= 0.18\, j_k \quad \underline{\text{(60\% leave nest, 30\% of those succeed)}} \\ a_{k+1} &= 0.71\, s_k + 0.94\, a_k \quad \text{(}\underline{\text{adults live about 20 yrs}}\text{)} \end{aligned}$$

The stage-matrix model is a difference equation of the form $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Such an equation is often called a **dynamical system** (or a **discrete linear dynamical system**) because it describes the changes in a system as time passes.

We have

$$\mathbf{x}_{k+1} = A\mathbf{x}_k = A(A\mathbf{x}_{k-1}) = A(A(A\mathbf{x}_{k-2})) = \ldots = A\,A\,A \ldots A\,A\,\mathbf{x}_0 = A^{k+1}\mathbf{x}_0$$

Then, the evolution is determined by the behavior of the powers of A.

A subject of interest to demographers is the movement of populations or groups of people from one region to another. The simple model here considers the changes in the population of a certain city and its surrounding suburbs over a period of years.

Fix an initial year—say, 2014—and denote the populations of the city and suburbs that year by r_0 and s_0, respectively. Let $\mathbf{x}_0$ be the population vector

$$\mathbf{x}_0 = \begin{bmatrix} r_0 \\ s_0 \end{bmatrix} \begin{array}{l} \text{City population, 2014} \\ \text{Suburban population, 2014} \end{array}$$

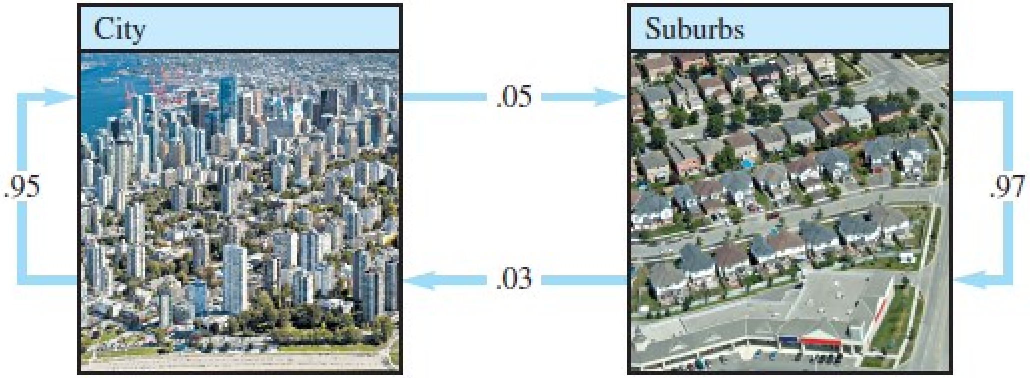
For 2015 and subsequent years, denote the populations of the city and suburbs by the vectors

$$\mathbf{x}_1 = \begin{bmatrix} r_1 \\ s_1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} r_2 \\ s_2 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} r_3 \\ s_3 \end{bmatrix}, \ldots$$

$\mathbf{x}_1 = M\mathbf{x}_0$ where M is the **migration matrix** determined by the following table:

$$\begin{array}{ccc} \text{From:} & & \\ \text{City} \quad \text{Suburbs} & & \text{To:} \\ \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} & & \begin{array}{l} \text{City} \\ \text{Suburbs} \end{array} \end{array}$$

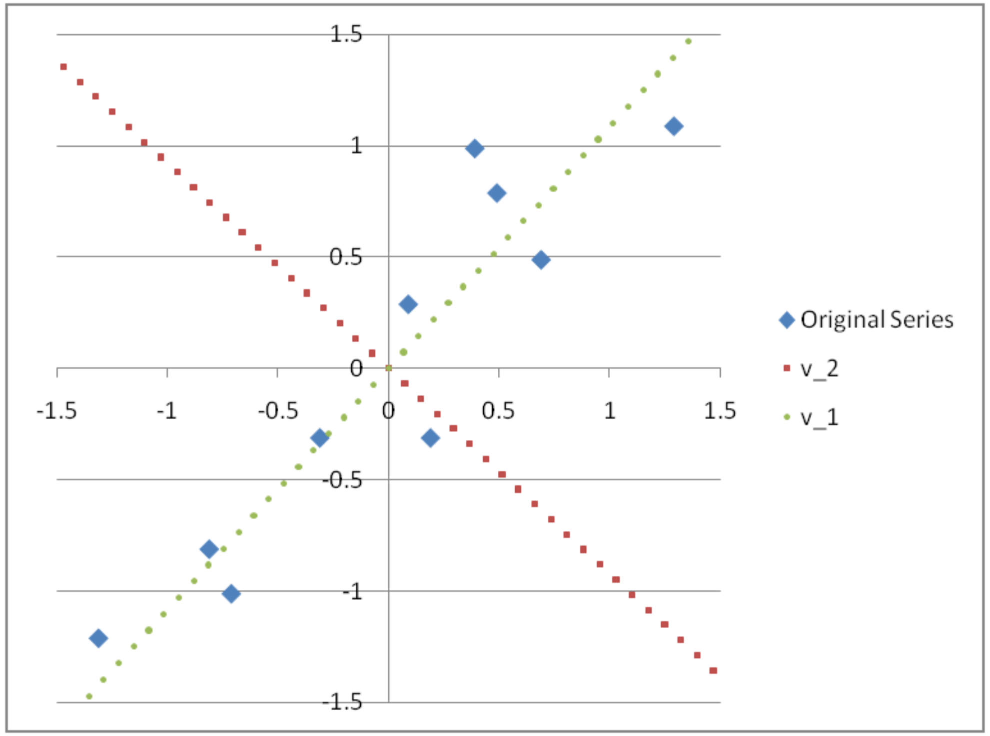
Mathematical model



Then $\mathbf{x}_k = M^k\mathbf{x}_0$

Principal Component Analysis (PCA)

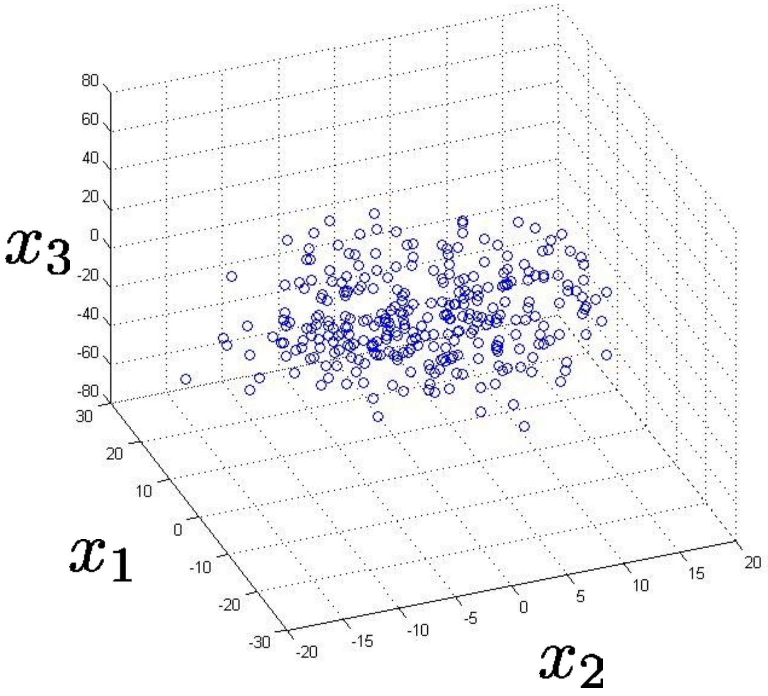
x	y
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3.0
2.3	2.7
2.0	1.6
1.0	1.1
1.5	1.6
1.1	0.9

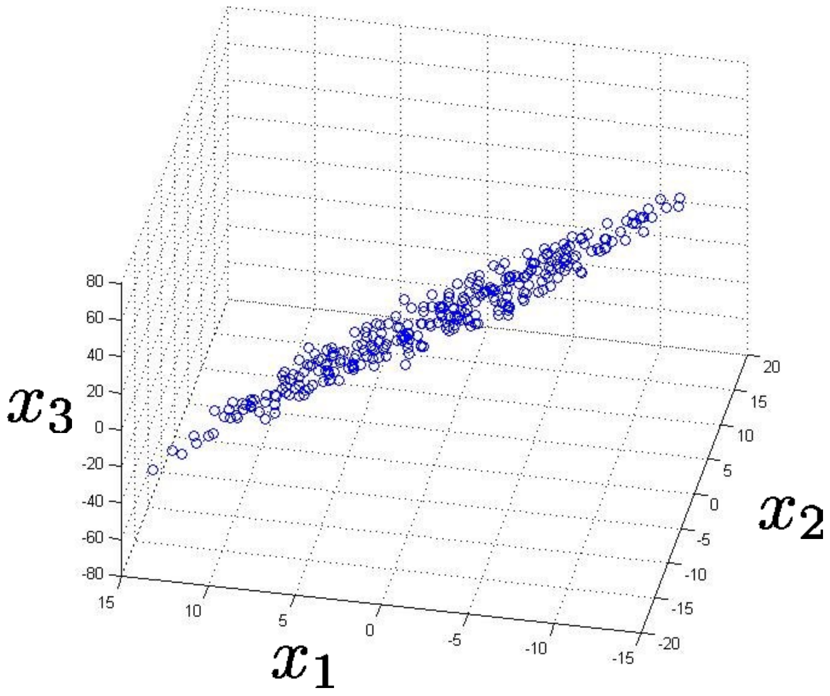


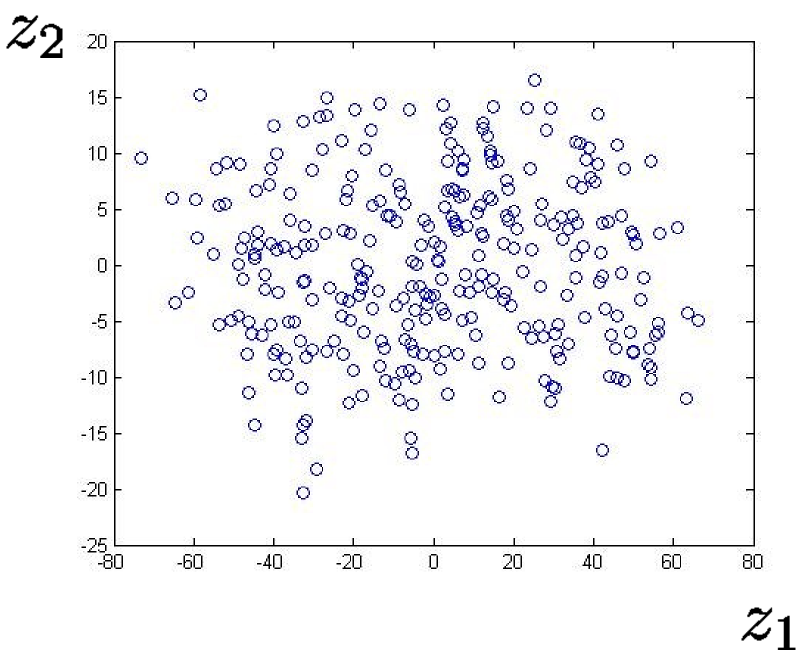
- PCA projects the data along the directions where the data varies most.
- These directions are determined by the eigenvectors of the covariance matrix corresponding to the largest eigenvalues.
- The magnitude of the eigenvalues corresponds to the variance of the data along the eigenvector directions.
- Find the projection that best preserves the variance.
- PCA preserves as much information as possible by minimizing the “reconstruction” error:

Data Compression

Example: Reduce data from 3D to 2D







Andrew Ng

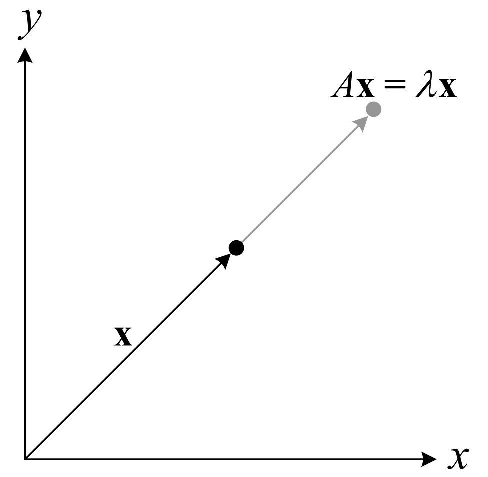
Eigenvalues and Eigenvectors

Almost all vectors change direction, when they are multiplied by A. Certain exceptional vectors **x** are in the same direction as A**x**. Those are the "**eigenvectors**". Multiply an eigenvector by A, and the vector A**x** is a number λ times the original **x**.

(The term eigenvalue is from the German word *Eigenwert,* meaning "proper value")

- Eigenvalue and Eigenvector:

A: an $n \times n$ matrix
λ: a scalar (could be **zero**)
$\mathbf{x}$: a **nonzero** vector in R^n



Geometric Interpretation

- Eigenvalue problem (one of the most important problems in the linear algebra):

If A is an $n \times n$ matrix, do there exist nonzero vectors $\mathbf{x}$ in R^n such that $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$?

Eigenvalue

$$\underset{\uparrow}{A\mathbf{x}} = \overset{\downarrow}{\lambda}\underset{\uparrow}{\mathbf{x}}$$

Eigenvector

Introduction to Eigenvalues and Eigenvectors

The **eigenvalue** tells whether the special vector $\mathbf{x}$ is stretched or shrunk or reversed or left unchanged—when it is multiplied by A. The eigenvalue could be zero! Then $A\mathbf{x} = 0\,\mathbf{x}$ means that this eigenvector $\mathbf{x}$ is in the nullspace (the space of the vectors such that $A\mathbf{x}=\mathbf{0}$). If A is the identity matrix, every vector has $A\mathbf{x} = \mathbf{x}$. All vectors are eigenvectors of I.

All eigenvalues "lambda" are =1. This is unusual to say the least. Most 2 by 2 matrices have *two* eigenvector directions and *two* eigenvalues. We will show that det $(A - \lambda I) = 0$:

$A\mathbf{x} = \lambda\,\mathbf{x}$ for $\mathbf{x} \neq \mathbf{0}$ and a scalar λ, so $(A - \lambda I)\mathbf{x} = \mathbf{0}$ and we have a non trivial solution if and only if det $(A - \lambda I) \neq 0$

Example Let $A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}$, $\det\begin{bmatrix} .8-\lambda & .3 \\ .2 & .7-\lambda \end{bmatrix} = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = (\lambda - 1)\left(\lambda - \frac{1}{2}\right)$ =0 then $\lambda=1$, and $\lambda=1/2$.

For those numbers, the matrix $(A - \lambda I)$ becomes *singular* (zero determinant). The eigenvectors **V**1 and **V**2 are in the nullspaces of $(A - I)$ and $(A - 1/2\, I)$.

$(A - I)\,\mathbf{V}1 = \mathbf{0} \;\Rightarrow\; \mathbf{V}_1 = t\begin{bmatrix} 1 \\ 2/3 \end{bmatrix}$ for any real value t

$(A - \tfrac{1}{2} I)\mathbf{V}2 = \mathbf{0} \;\Rightarrow\; \mathbf{V}_2 = t\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ for any real value t

Example Verifying eigenvalues and eigenvectors

$$A=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix} \quad \mathbf{x}_1=\begin{bmatrix}1\\ 0\end{bmatrix} \quad \mathbf{x}_2=\begin{bmatrix}0\\ 1\end{bmatrix}$$

Eigenvalue

$$A\mathbf{x}_1=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}=\begin{bmatrix}2\\ 0\end{bmatrix}=2\begin{bmatrix}1\\ 0\end{bmatrix}=2\mathbf{x}_1$$

Eigenvector

In fact, for each eigenvalue, it has infinitely many eigenvectors. For example, for $\lambda = 2$, $[3\ 0]^T$ or $[5\ 0]^T$ are both corresponding eigenvectors. Moreover, $([3\ 0] + [5\ 0])^T$ is still an eigenvector.

Eigenvalue

$$A\mathbf{x}_2=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix}=\begin{bmatrix}0\\ -1\end{bmatrix}=-1\begin{bmatrix}0\\ 1\end{bmatrix}=(-1)\mathbf{x}_2$$

Eigenvector

Summary To solve the eigenvalue problem for an n by n matrix, follow these steps:

1. ***Compute the determinant of*** $A - \lambda I$. With λ subtracted along the diagonal, this determinant starts with λ^n or $-\lambda^n$. It is a polynomial in λ of degree n.
2. ***Find the roots of this polynomial***, by solving $\det(A - \lambda I) = 0$. The n roots are the n eigenvalues of A. They make $A - \lambda I$ singular.
3. For each eigenvalue λ, ***solve*** $(A - \lambda I)x = 0$ ***to find an eigenvector*** x.

Theorem (The eigenspace of A corresponding to λ)
If A is an $n \times n$ matrix with an eigenvalue λ, then the set of all eigenvectors of λ together with the zero vector is a subspace of R^n. This subspace is called the eigenspace of λ.

Remark Are eigenvectors are unique? If $\mathbf{x}$ is an eigenvector, then $\beta\mathbf{x}$ is also an eigenvector and $\beta\lambda$ is an eigenvalue

$$\mathbf{A}(\beta\mathbf{x}) = \beta(\mathbf{Ax}) = \beta(\lambda\mathbf{x}) = \lambda(\beta\mathbf{x})$$

Remark The eigenvalues (roots of a polynomial) could be complex numbers!

Calculating the Eigenvectors/values

- Expand the det(**A** - λ**I**) = 0 for a 2 x 2 matrix

$$\det(A-\lambda I)=\det\left(\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}-\lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right)=0$$

$$\det\begin{bmatrix}a_{11}-\lambda & a_{12}\\ a_{21} & a_{22}-\lambda\end{bmatrix}=0\Rightarrow(a_{11}-\lambda)(a_{22}-\lambda)-a_{12}a_{21}=0$$

$$\lambda^2-\lambda(a_{11}+a_{22})+(a_{11}a_{22}-a_{12}a_{21})=0$$

- For a 2 x2 matrix, this is a simple quadratic equation with two solutions (maybe complex)

$$\lambda=(a_{11}+a_{22})\pm\sqrt{\frac{(a_{11}+a_{22})^2}{4(a_{11}a_{22}-a_{12}a_{21})}}$$

- This "characteristic equation" can be used to solve for **x**

Eigenvalue example

- Consider,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \Rightarrow \begin{cases} \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0 \\ \lambda^2 - (1+4)\lambda + (1\cdot 4 - 2\cdot 2) = 0 \\ \lambda^2 = (1+4)\lambda \Rightarrow \lambda = 0, \lambda = 5 \end{cases}$$

- The corresponding eigenvectors can be computed as

$$\lambda = 0 \Rightarrow \left[\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right] \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1x + 2y \\ 2x + 4y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\lambda = 5 \Rightarrow \left[\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} - \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}\right] \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -4x + 2y \\ 2x - 1y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

  - For $\lambda = 0$, one possible solution is $\mathbf{x} = (2, -1)$
  - For $\lambda = 5$, one possible solution is $\mathbf{x} = (1, 2)$

Warning: we compute eigenvalues using the determinant only for very low dimensional case as an exercise, in the applications we must to consider efficient numerical methods and we (usually) have to calculate only a few eigenvalues / eigenvectors.

Theorem

If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix A, then the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.

Let A be a square n x n matrix with *n* linearly independent eigenvectors (a "non-defective" matrix)
Let P have the eigenvectors as columns:

$$P = [\,V_1 \;\; V_2 \;\; \ldots \;\; V_n]$$

Then, AP can be written

$$\mathrm{AP} = A[\mathbf{V}_1 \;\; \mathbf{V}_2 \;\; \ldots \;\; \mathbf{V}_\mathrm{n}] = [\lambda_1\mathbf{V}_1 \;\; \lambda_2\mathbf{V}_2 \;\; \ldots \;\; \lambda_n\mathbf{V}_\mathrm{n}] = [\mathbf{V}_1 \;\; \mathbf{V}_2 \;\; \ldots \;\; \mathbf{V}_\mathrm{n}] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ldots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} = [\mathbf{V}_1 \;\; \mathbf{V}_2 \;\; \ldots \;\; \mathbf{V}_\mathrm{n}]\,\Lambda$$

Thus $AP=P\Lambda$, or $P^{-1}AP= \Lambda$, And $A=P\Lambda P^{-1}$ with Λ diagonal matrix with eigenvalues.

A is called diagonalizable

Example Eigenvalue problems and diagonalization

$$A = \begin{bmatrix} 1 & 3 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Characteristic equation:

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -3 & 0 \\ -3 & \lambda - 1 & 0 \\ 0 & 0 & \lambda + 2 \end{vmatrix} = (\lambda - 4)(\lambda + 2)^2 = 0$$

$$\text{The eigenvalues}: \lambda_1 = 4,\ \lambda_2 = -2,\ \lambda_3 = -2$$

$$(1)\ \lambda = 4 \Rightarrow \text{ the eigenvector } \mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

(**one possible eigenvector**)

(2) $\lambda = -2 \Rightarrow$ the eigenvector

$$\mathbf{p}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{p}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3] = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } P^{-1}AP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

($\mathbf{p}_1$ $\mathbf{p}_2$ $\mathbf{p}_3$ are linear independent)

NOTE If $P = [\mathbf{p}_2 \quad \mathbf{p}_1 \quad \mathbf{p}_3]$

$$= \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \Rightarrow P^{-1}AP = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Compute the power of A: a simple case.

Let A a diagonalizable matrix: $A=P\Lambda P^{-1}$ for a suitable P. Then

$A^k = P\,\Lambda P^{-1} P\,\Lambda\, P^{-1}\ \ldots\ P\Lambda\, P^{-1} = P\,\Lambda^k P^{-1}$

And the long range behavior is determined by the power of the eigenvalues.

Example Let initial population $\mathbf{x}_0 = \begin{bmatrix} 8000 \\ 2000 \end{bmatrix}$ and a transition matrix $A = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$

$\mathbf{x}_{k+1} = A\,\mathbf{x}_k \quad k=0,1,2,\ldots$

Eigenvalues of A: $\lambda_1 = 1$ (easy from the entries of the matrix A… why?) $\lambda_2 = 1/2$ For the eigenvectors we choose:

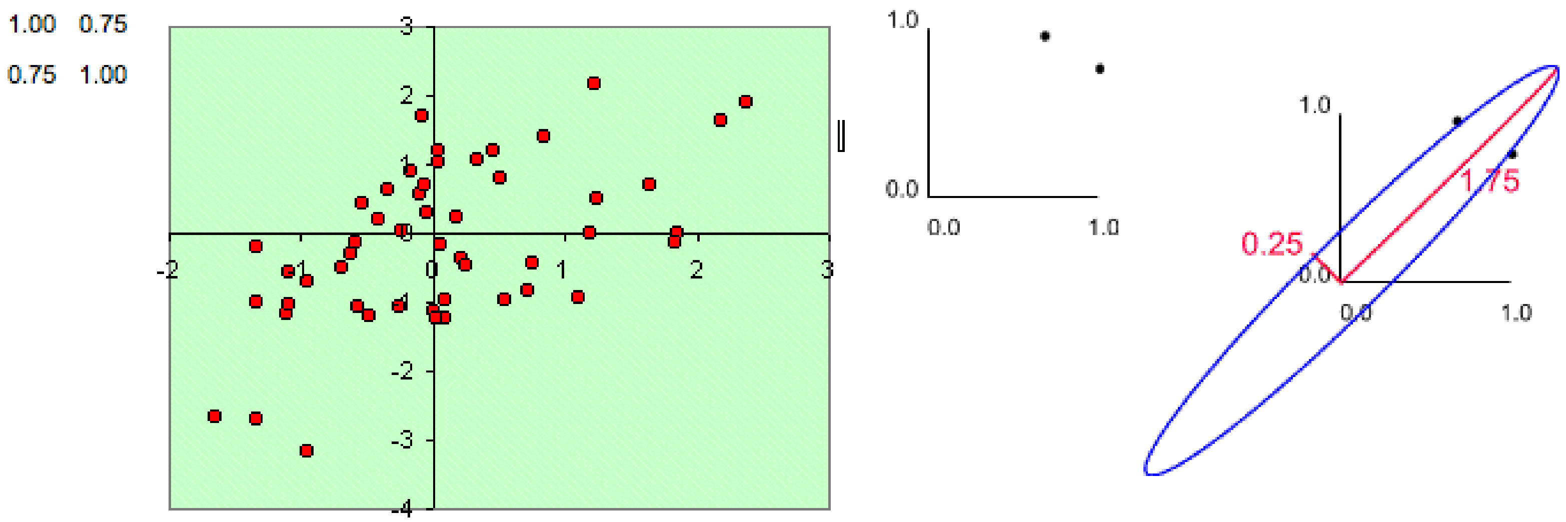
$\mathbf{V}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\mathbf{V}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{x}_k = A^k\,\mathbf{x}_0 = P\,\Lambda^k P^{-1}\,\mathbf{x}_0$ where $P=[\mathbf{V}_1\ \mathbf{V}_2\,]$ then (after some computation)

$\mathbf{x}_k = 2000\,\mathbf{V}_1 - 4000\left(\frac{1}{2}\right)^k \mathbf{V}_2$ when $k>>1$ $\mathbf{x}_k \sim 2000\,\mathbf{V}_1$

Physical interpretation

- Consider a covariance matrix, **A**, i.e., A = 1/n S S^T for some S (two variable, n subjects)

$$\mathbf{A}=\begin{bmatrix}1 & .75\\ .75 & 1\end{bmatrix}\Rightarrow \lambda_1=1.75, \lambda_2=0.25$$



- Error ellipse with the major axis as the larger eigenvalue and the minor axis as the smaller eigenvalue

Physical interpretation

- First principal component is the direction of greatest variability (covariance) in the data
- Second is the next orthogonal (uncorrelated) direction of greatest variability
  - So first remove all the variability along the first component, and then find the next direction of greatest variability
- And so on …
- Thus each eigenvectors provides the directions of data variances in decreasing order of eigenvalues

Principal Component Analysis

- Principal component analysis
  - It is a way of identifying the underlying patterns in data
  - It can extract information in a large data set with many variables and approximate this data set with fewer factors
  - In other words, it can reduce the number of variables to a more manageable set

- Steps of the principal component analysis

- Step 1: Get some data
- Step 2: Subtract the mean
- Step 3: Calculate the covariance matrix
- Step 4: Calculate the eigenvectors and eigenvalues of the covariance matrix
- Step 5: Deriving the transformed data set
- Step 6: Getting the original data back

Step 1:

x	y
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3.0
2.3	2.7
2.0	1.6
1.0	1.1
1.5	1.6
1.1	0.9
[redacted]	

Step 2:

x	y
0.69	0.49
-1.31	-1.21
0.39	0.99
0.09	0.29
1.29	1.09
0.49	0.79
0.19	-0.31
-0.81	-0.81
-0.31	-0.31
-0.71	-1.01
[redacted]	

$$\equiv X^T = \begin{bmatrix} x & y \end{bmatrix}$$

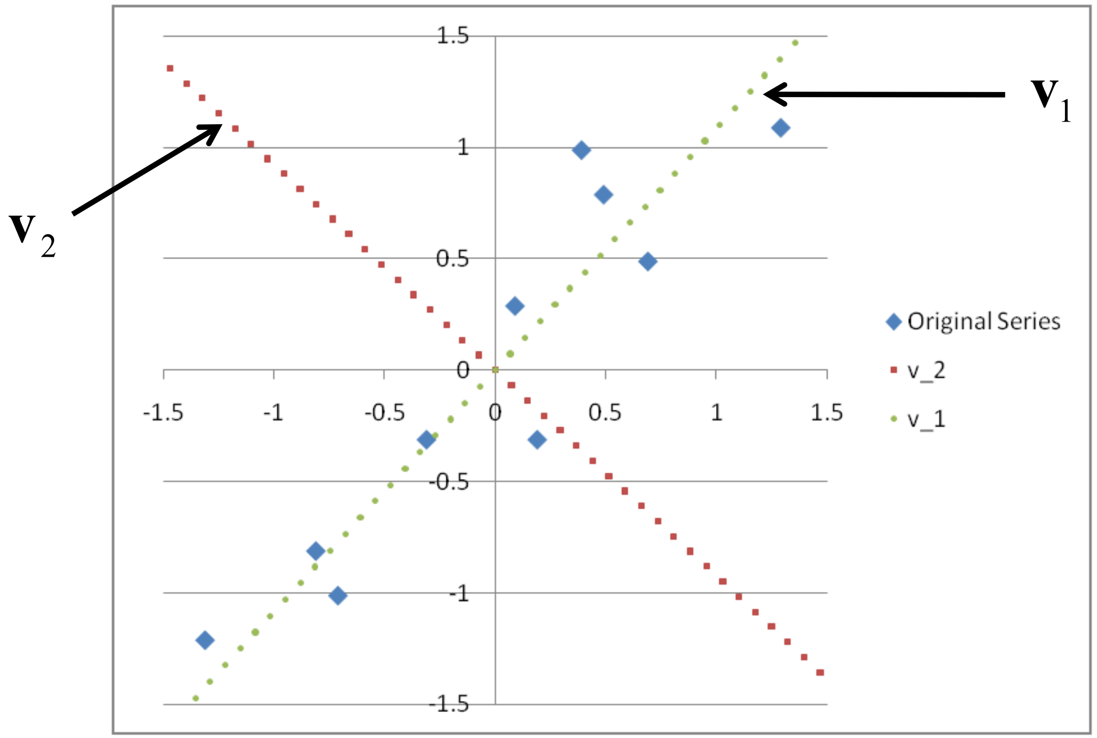
Step 3:

$$\mathrm{var}(X^T) = E\left[XX^T\right] = E\left[\begin{bmatrix} x^T \\ y^T \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix}\right] = E\begin{bmatrix} x^T x & x^T y \\ y^T x & y^T y \end{bmatrix}$$

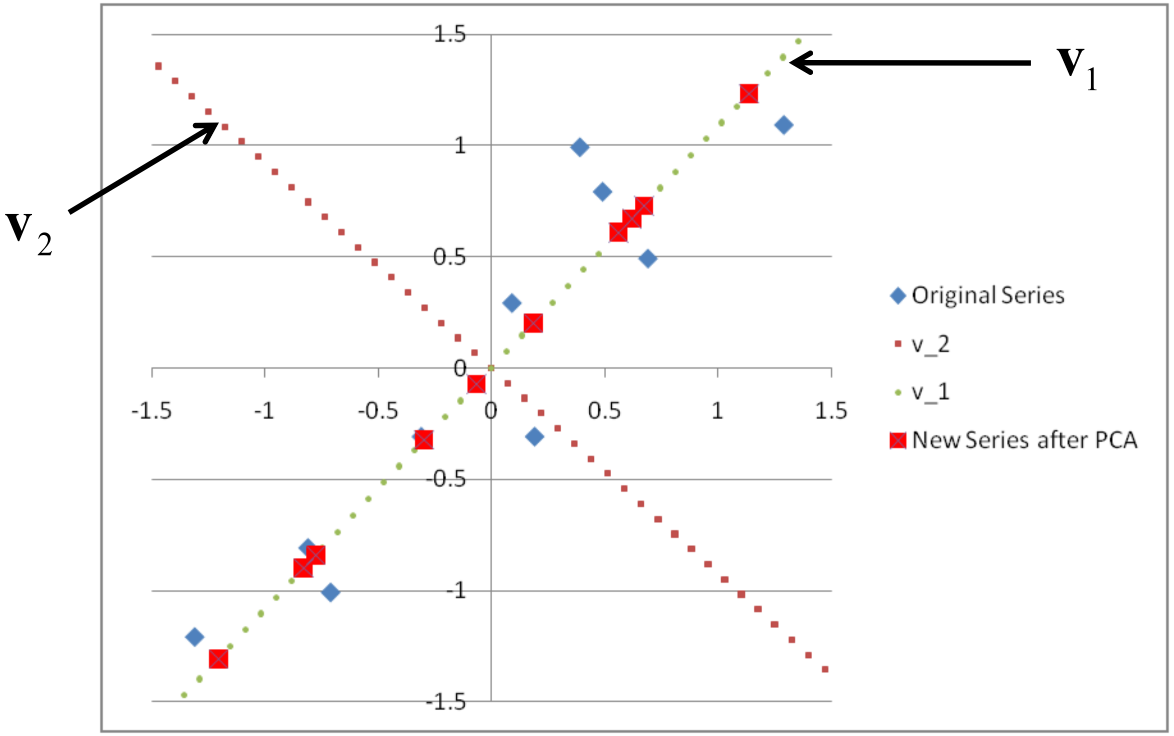
$$= \begin{bmatrix} \mathrm{var}(x) & \mathrm{cov}(x,y) \\ \mathrm{cov}(x,y) & \mathrm{var}(y) \end{bmatrix} = \begin{pmatrix} 0.616556 & 0.615444 \\ 0.615444 & 0.716556 \end{pmatrix} \equiv A$$

- **Step 4: Calculate the eigenvectors and eigenvalues of the covariance matrix *A***

$$\lambda_1 = 1.284028,\ \mathbf{v}_1 = \begin{pmatrix} -0.67787 \\ -0.73518 \end{pmatrix} \qquad \lambda_2 = 0.049083,\ \mathbf{v}_2 = \begin{pmatrix} -0.73518 \\ 0.67787 \end{pmatrix}$$



1. The two eigenvectors are orthogonal to each other-
2. $\mathbf{v}_1$ eigenvector (corresponding to the largest eigenvalue λ_1) is just like a best-fit regression line
3. $\mathbf{v}_2$ seems less important to explain the data since the projection of each node on the $\mathbf{v}_2$ axis is very close to zero
4. The interpretation of $\mathbf{v}_1$ is the new axis which retains as much as possible the variance information that was contained in the original two dimensions



- If only the principal component *x'* is considered in the Principal Component Analysis (PCA), it is equivalent to project all points onto the $\mathbf{v}_1$ vector
- It can be observed in the above figure that the projection onto $\mathbf{v}_1$ vector can retains as much as possible the "interpoint" distance information (variance) that was contained in the original series of (*x*, *y*)