Math 2270-002 Week 15 notes
We will not necessarily finish the material from a given day's notes on that day. We may also add or subtract some material as the week progresses, but these notes represent an in-depth outline of what we plan to cover. These notes cover material in 7.1-7.2, with forays into applications and latter sections of Chapter 7

Mon Dec 3

- 7.1-7.2 Diagonalizing quadratic forms and surfaces and curves defined implicitly with quadratic equations, via the spectral theorem continued;   spectral decomposition Theorem (section 7.1) - with hints of Professor Tom Alberts' presentation tomorrow on "Principal Component Analysis" (PCA), and its uses with dimension reduction ,for "covariance matrices" of large data sets. (See e.g. Wikipedia.) Professor Alberts' lecture will tie in to a ground-breaking 2008 paper in "Nature", which foreshadows the subsequent rise of genetic geneology, i.e. the methods with which one can estimate where on earth an individual's long ago ancestors lived.

"Genes Mirror Geography in Europe"

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2735096/

Announcements:

In the last week of the course!

Warm-up Exercise: Today we learn a new way to multiply matrices. And it's useful

a) Compute

$$\underset{A}{\begin{bmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 2 \end{bmatrix}} \underset{B}{\begin{bmatrix} 2 & -3 & 0 \\ 1 & 0 & 4 \end{bmatrix}} = \begin{bmatrix} 1 & -3 & -4 \\ 4 & -6 & 0 \\ 2 & 0 & 8 \end{bmatrix}$$

$\text{entry}_{ij}(AB) = \text{row}_i(A) \cdot \text{column}_j(B)$ "inner products"

b) Then - using the columns of A, and corresponding rows of B (!!), compute the sum

"outer products"

$$\underbrace{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}}_{col_1(A)} \underbrace{\begin{bmatrix} 2 & -3 & 0 \end{bmatrix}}_{row_1(B)} + \underbrace{\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 4 \end{bmatrix}}_{col_2(A)\, row_2(B)} = $$

$$: \sum_{p=1}^{2} col_p(A)\, row_p(B)$$

$$= \begin{bmatrix} 2 & -3 & 0 \\ 4 & -6 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 4 \\ 0 & 0 & 0 \\ 2 & 0 & 8 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 4 \\ 4 & -6 & 0 \\ 2 & 0 & 8 \end{bmatrix}$$

Recall from Friday ...

<u>Spectral Theorem</u> Let A be an $n \times n$ symmetric matrix. Then all of the eigenvalues of A are real, and there exists an orthonormal eigenbasis $B = \{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \ldots \underline{\boldsymbol{u}}_n\}$ consisting of eigenvectors for A. Eigenspaces with different eigenvalues are automatically orthogonal to each other. If any eigenspace has dimension greater than 1, its orthonormal basis may be constructed via *Gram − Schmidt*. (We proved parts of the spectral theorem on Friday.)

One application, as discussed on Friday:
<u>Diagonalization of quadratic forms</u>: Let

$$Q(\underline{\mathbf{x}}) = \sum_{i,j=1}^{n} a_{ij} x_i x_j = \underline{\mathbf{x}}^T A \underline{\mathbf{x}}$$

for a symmetric matrix A, with real entries. A symmetric $\Rightarrow$ by the spectral theorem there exists an orthonormal eigenbasis $B = \{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \ldots \underline{\boldsymbol{u}}_n\}$.

For the corresponding orthogonal matrix

$$P = [\underline{\boldsymbol{u}}_1 | \underline{\boldsymbol{u}}_2 | \cdots | \underline{\boldsymbol{u}}_n]$$

$$A P = P D$$
$$D = P^T A P,$$

where D is the diagonal matrix of eigenvalues corresponding to the eigenvectors in P. So for the change of variables

$$\underline{\mathbf{x}} = P \underline{\mathbf{y}}$$

where $\underline{\mathbf{y}} = [\underline{\mathbf{x}}]_B$ and $P = P_{E \leftarrow B}$, we have

$$Q(\underline{\mathbf{x}}) = \underline{\mathbf{x}}^T A \underline{\mathbf{x}}$$

$$= \underline{\mathbf{y}}^T P^T A P \underline{\mathbf{y}} = \underline{\mathbf{y}}^T D \ \underline{\mathbf{y}}$$

$$= \sum_{i=1}^{n} \lambda_i y_i^2 .$$

So by the orthogonal change of variables all cross terms have been removed. Applications include conic curves, quartic surfaces, multivariable second derivative test, singular value decomposition theorem, and more.

Example Identify and sketch the surface defined implicitly by

typo

$$x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_2 - 4x_1x_3 - 4x_2x_3 = 8.$$

Exercise 1) Find the symmetric matrix so that

typo

$$x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_2 - 4x_1x_3 - 4x_2x_3 = \underline{\boldsymbol{x}^T A\,\boldsymbol{x}}.$$

Recall that

$$\underline{\boldsymbol{x}}^T A\,\boldsymbol{x} = \sum_{i,j=1}^{n} a_{ij}x_i x_j .$$

$a_{12}x_1x_2 + a_{21}x_2x_1 = -2x_1x_2$

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 1 & -1 & -2 \\ -1 & 1 & -2 \\ -2 & -2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

If we found the matrix correctly technology tells us that

$$E_{\lambda=-2} = span\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\},\ E_{\lambda=2} = span\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\},\ E_{\lambda=4} = span\left\{\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}\right\}.$$

(positively oriented in this order)

$\perp$ $\perp$

$\perp$

$$x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_2 - 4x_1x_3 - 4x_2x_3 = 8$$

$$\underline{\boldsymbol{x}}^T A\, \underline{\boldsymbol{x}} = 8$$

For

$$P = \begin{array}{c} \begin{array}{ccc} \lambda=-2 & \lambda=2 & \lambda=4 \end{array} \\ \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix} \end{array}$$

$P^{-1} = P^T$

$AP = PD$

$P^TAP = D.$

$$P^T A\, P = D$$

original quadratic form → $\underline{\boldsymbol{x}}^T A\, \underline{\boldsymbol{x}} = 8$

$$\underline{\boldsymbol{x}} = P\, \underline{\boldsymbol{y}}$$

$P = [P]_{E \leftarrow B}$

$\vec{y} = [\vec{x}]_B$

$$\underline{\boldsymbol{y}}^T P^T A\, P\, \underline{\boldsymbol{y}} = 8$$

$$\underline{\boldsymbol{y}}^T \mathrm{D}\, \underline{\boldsymbol{y}} = 8$$

$$-2y_1^2 + 2y_2^2 + 4y_3^2 = 8.$$

One can try to sketch this in the rotated coordinate system. It's an elliptic hyperboloid of one sheet. :-)
<u>Note:</u> You only need to know the eigenvalues to sketch the surface in the new coordinate system.

$$2y_2^2 + 4y_3^2 = 8 + 2y_1^2$$

$$y_2^2 + 2y_3^2 = 4 + y_1^2$$

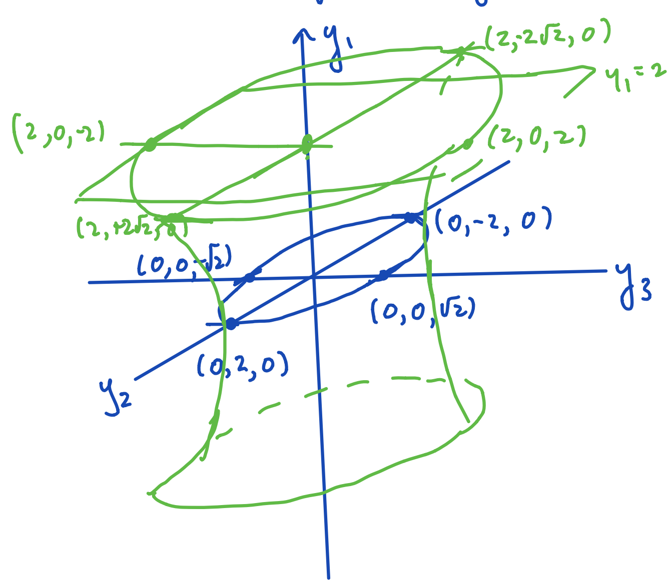
traces in planes with $y_1 = \text{const}$ are ellipses.

$y_1 = 2$

$y_2^2 + 2y_3^2 = 8$

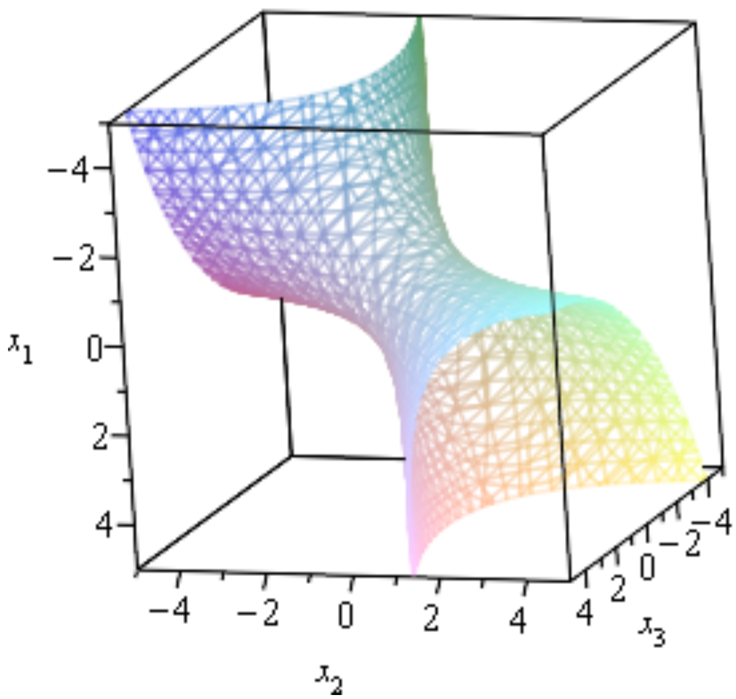
$y_1 = 0$

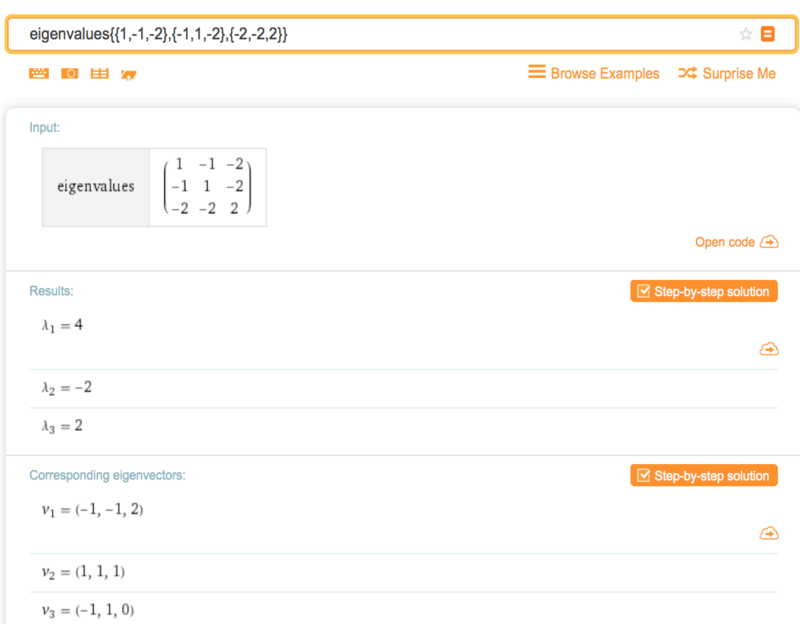
$\to y_2^2 + 2y_3^2 = 4$



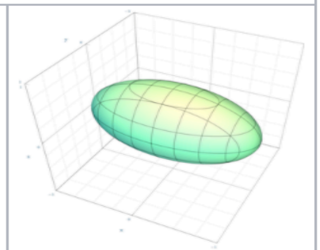
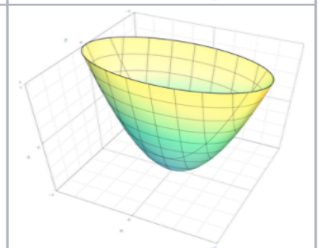
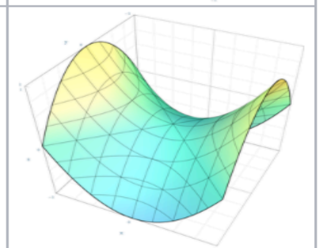
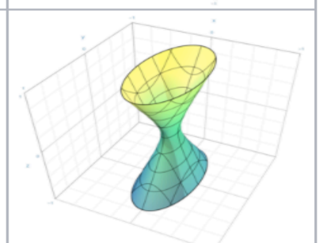
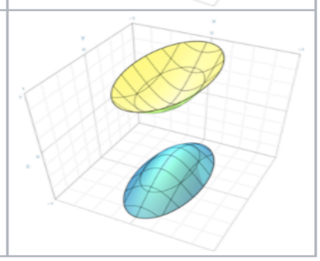
In the original coordinate system we see indications of the rotated one:

> $with(plots):$
$implicitplot3d\left(x_1^2 + x_2^2 + 2 \cdot x_3^2 - 2 \cdot x_1 \cdot x_2 - 4 \cdot x_1 \cdot x_2 - 4 \cdot x_2 x_3 = 8, x_1 = -5 \,..5, x_2 = -5 \,..5, x_3 = -5 \,..5, grid = [20, 20, 20]\right);$





from Wikipedia, "quadric surfaces". There is also a Wikipedia page on conic sections.

Non-degenerate real quadric surfaces		
Ellipsoid	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$	
Elliptic paraboloid	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - z = 0$	
Hyperbolic paraboloid	$\frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0$	
Elliptic hyperboloid of one sheet	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$	
Elliptic hyperboloid of two sheets	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$	

Material we need for Prof. Alberts' guest lecture Tuesday on Principal Component Analysis. (The text discusses some of this background material in 7.1, 7.2. The text also has a section on PCA in Chapter 7, although their approach reads differently than the one we will take.)

Definition: The quadratic form $Q(\underline{x}) = \sum_{i,j=1}^{n} a_{ij} x_i x_j = \underline{x}^T A \underline{x}$ (for A a symmetric matrix) is called *positive definite* if

$$Q(\underline{x}) > 0 \quad \text{for all } \underline{x} \neq \underline{0}.$$

$\vec{x}^T A \vec{x}$
$= \vec{y}^T D \vec{y}$
$= \sum_{i=1}^{n} \lambda_i y_i^2$

From the preceding discussion, we see that this is the same as saying that all of the eigenvalues of A are positive. If

$$Q(\underline{x}) \geq 0 \quad \text{for all } \underline{x} \neq \underline{0}$$

then Q is called *positive semi-definite.* In terms of eigenvalues, this is equivalent to saying all of the eigenvalues of A are non-negative.

Two important examples of positive semi-definite matrices from data analysis: Let $\{\underline{x}_1, \underline{x}_2, \ldots \underline{x}_n\}$ be n vectors of data in $\mathbb{R}^N$. (For example, $\underline{x}_1$ could be the vector of "log-heights" from our previous discussion, and $\underline{x}_2$ could be the corresponding vector of "log-weights") "Normalize" each vector so that the average of its entries is zero. In other words, replace each $\underline{x}_i$ with

$$\underline{z}_i = \underline{x}_i - m_i \underline{1}$$

where m_i is the mean of the entries in $\underline{x}_i$ and $\underline{1}$ is a vector of $1's$.

Definition: The *covariance* matrix A for the data set $\{\underline{x}_1, \underline{x}_2, \ldots \underline{x}_n\}$ is the symmetric matrix with $a_{ij} = \underline{z}_i \bullet \underline{z}_j$, i.e.

$$A = \begin{bmatrix} - - - \underline{z}_1^T - - - \\ - - - \underline{z}_2^T - - - \\ \vdots \\ - - - \underline{z}_n^T - - - \end{bmatrix} \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ \underline{z}_1 & \underline{z}_2 & \cdots & \underline{z}_n \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} := Z^T Z$$

Note, $\underline{z}_i \bullet \underline{z}_j = \|\underline{z}_i\| \, \|\underline{z}_j\| \cos\theta_{ij}$ where θ_{ij} is the angle between the vectors $\underline{z}_i, \underline{z}_j$ so the entries in this matrix quantify how the two vectors "co-vary" or not. The diagonal entries, suitably normalized, measure the "variance" of the corresponding $\underline{z}_i$. The matrix A is positive semi-definite because the corresponding quadratic form satisfies

$$Q(\underline{c}) = \sum_{i,j=1}^{n} a_{ij} c_i c_j = \underline{c}^T A \underline{c} = \underline{c}^T Z^T Z \underline{c} = (Z\underline{c})^T (Z\underline{c}) = \|Z\underline{c}\|^2 \geq 0.$$

The *correlation* matrix R for the same data set is the matrix with

$$r_{ij} = \frac{\underset{\sim}{z}_i}{\|\underset{\sim}{z}_i\|} \bullet \frac{\underset{\sim}{z}_j}{\|\underset{\sim}{z}_j\|}$$

being the *correlation coefficient* $\cos\left(\theta_{ij}\right)$ between $\underset{\sim}{z}_i$ and $\underset{\sim}{z}_j$.

$$R = \begin{bmatrix} --- & \dfrac{\underset{\sim}{z}_1^T}{\|\underset{\sim}{z}_1^T\|} & --- \\ --- & \dfrac{\underset{\sim}{z}_2^T}{\|\underset{\sim}{z}_2^T\|} & --- \\ & \vdots & \\ --- & \dfrac{\underset{\sim}{z}_n^T}{\|\underset{\sim}{z}_n^T\|} & --- \end{bmatrix} \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ \dfrac{\underset{\sim}{z}_1}{\|\underset{\sim}{z}_1\|} & \dfrac{\underset{\sim}{z}_2}{\|\underset{\sim}{z}_2\|} & \cdots & \dfrac{\underset{\sim}{z}_n}{\|\underset{\sim}{z}_n\|} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

An analogous computation to the one for the covariance matrix shows that R is positive semidefinite.

Theorem: The "outer product" way of computing the matrix product $A\,B$. (Section 2.4 topic on partitioned matrices that we skipped at the time....our usual way is with dot product or rows of A with columns of B, aka an "inner product").

(1) first, notice that the product of an $m \times 1$ column vector with a $1 \times n$ row vector is an $m \times n$ matrix; and the entry in the i^{th} row and j^{th} column of the product is the product of the i^{th} entry in the column vector with the j^{th} entry in the row vector. For example:

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} b_1 & b_2 \end{bmatrix} = \begin{bmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \\ a_3 b_1 & a_3 b_2 \end{bmatrix}.$$

→ the ij entry is just $a_i b_j$

(2) Let $A_{m \times p}$ and $B_{p \times n}$. Express A in terms of its columns, and B in terms of its rows:

$$A = \begin{bmatrix} | & | & & | \\ \underline{\boldsymbol{a}}_1 & \underline{\boldsymbol{a}}_2 & \cdots & \underline{\boldsymbol{a}}_p \\ | & | & & | \\ | & | & & | \end{bmatrix} \qquad B = \begin{bmatrix} ---\underline{\boldsymbol{b}}_1--- \\ ---\underline{\boldsymbol{b}}_2--- \\ \vdots \\ ---\underline{\boldsymbol{b}}_p--- \end{bmatrix}.$$

Then

$$A\,B = \sum_{k=1}^{p} \underline{\boldsymbol{a}}_k\, \underline{\boldsymbol{b}}_k\,.$$

proof:
Using our usual ("inner product") way of computing matrix products,

$$entry_{i\,j}\,A\,B = row_i(A) \bullet col_j(B) = \sum_{k=1}^{p} a_{i\,k}\,b_{k\,j}.$$

Compare this to the outer product formula for the ij entry:

$$entry_{i\,j}\left(\sum_{k=1}^{p} \underline{\boldsymbol{a}}_k\, \underline{\boldsymbol{b}}_k\right) = \sum_{k=1}^{p} entry_{i\,j}\left(\underline{\boldsymbol{a}}_k\, \underline{\boldsymbol{b}}_k\right) = \sum_{k=1}^{p} entry_i\left(\underline{\boldsymbol{a}}_k\right) entry_j\left(\underline{\boldsymbol{b}}_k\right) = \sum_{k=1}^{p} a_{i\,k}\,b_{k\,j}\,.$$

SAME!!

Spectral decomposition for symmetric matrices. Let $A_{n \times n}$ be symmetric (and positive semi-definite, for the applications Prof. Alberts will talk about tomorrow). Order the eigenvalues as

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$$

and let

$$\{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \cdots, \underline{\boldsymbol{u}}_n\}$$

be corresponding orthonormal eigenvectors in $\mathbb{R}^n$. Let P be the orthogonal matrix

$$P = [\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \cdots \underline{\boldsymbol{u}}_n]$$

with

$$A\,P = P\,D$$

where D is the diagonal matrix with diagonal entries $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n$.

Then

$$A = P\,D\,P^T$$

$$= \begin{bmatrix} | & | & & | \\ \underline{\boldsymbol{u}}_1 & \underline{\boldsymbol{u}}_2 & \ldots & \underline{\boldsymbol{u}}_n \\ | & | & & | \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ : & \ldots & \ldots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} \begin{bmatrix} ---\underline{\boldsymbol{u}}_1^T--- \\ ---\underline{\boldsymbol{u}}_2^T--- \\ : \\ ---\underline{\boldsymbol{u}}_n^T--- \end{bmatrix}$$

$$= \begin{bmatrix} | & | & & | \\ \lambda_1 \underline{\boldsymbol{u}}_1 & \lambda_2 \underline{\boldsymbol{u}}_2 & \ldots & \lambda_n \underline{\boldsymbol{u}}_n \\ | & | & & | \\ | & | & & | \end{bmatrix} \begin{bmatrix} ---\underline{\boldsymbol{u}}_1^T--- \\ ---\underline{\boldsymbol{u}}_2^T--- \\ : \\ ---\underline{\boldsymbol{u}}_n^T--- \end{bmatrix}$$

$$\boxed{A = \lambda_1 \underline{\boldsymbol{u}}_1 \underline{\boldsymbol{u}}_1^T + \lambda_2 \underline{\boldsymbol{u}}_2 \underline{\boldsymbol{u}}_2^T + \ldots + \lambda_n \underline{\boldsymbol{u}}_n \underline{\boldsymbol{u}}_n^T}$$

This is the spectral decomposition of A. "Principal component analysis" of covariance or correlation matrices makes use of the fact that if only a few of the eigenvalues of A are large and the rest are near zero, then the corresponding leading terms in the expression above are a good approximation for the matrix A. And projection of the data onto the span of the corresponding important eigenvectors is a good facsimile of the original data, potentially in a much lower dimensional space.

Testing spectral decomposition in a small example from last week (like one of your homework exercises).

$$A = \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix}$$

$$E_{\lambda = \frac{9}{2}} = span\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\} \quad E_{\lambda = -\frac{1}{2}} = span\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$$

$$\underline{\boldsymbol{u}}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \underline{\boldsymbol{u}}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\lambda_1 \underline{\boldsymbol{u}}_1 \underline{\boldsymbol{u}}_1^T + \lambda_2 \underline{\boldsymbol{u}}_2 \underline{\boldsymbol{u}}_2^T = \frac{9}{2} \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \end{bmatrix} - \frac{1}{2} \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \end{bmatrix}$$

$$= \frac{9}{4}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & \frac{5}{2} \\ \frac{5}{2} & 2 \end{bmatrix} !!!!$$

attribute ~ like heights, weights, etc. (for 5 individuals)

Example with a covariance matrix. Consider the following four data vectors in $\mathbb{R}^5$. Note that each data vector already has mean equal to zero. What is the effective "dimension" of this data set, and in which directions is the "variance" of the data the greatest? (This is a very small example of "big data". :-))

$$\left\{ \begin{bmatrix} -2 \\ 0 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \\ 0 \end{bmatrix} \right\}.$$

a) Verify that the covariance matrix is given by

$$A = \begin{bmatrix} 8 & 0 & 0 & -8 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ -8 & 0 & 0 & 8 \end{bmatrix}.$$

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

b) Verify by observation the following orthogonal eigendata for A:

$$E_{\lambda=16} = span\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \right\}, \quad E_{\lambda=4} = span\left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \right\}, \quad E_{\lambda=0} = span\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

$$\begin{bmatrix} 8 & 0 & 0 & -8 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ -8 & 0 & 0 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ -16 & 0 & 0 & 0 \end{bmatrix} \checkmark$$

$$\text{original data vectors: } \left\{ \begin{bmatrix} -2 \\ 0 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \\ 0 \end{bmatrix} \right\}.$$

$$\text{covariance matrix } A = \begin{bmatrix} 8 & 0 & 0 & -8 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ -8 & 0 & 0 & 8 \end{bmatrix}$$

$$E_{\lambda=16} = span\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \right\},\ E_{\lambda=4} = span\left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \right\},\ E_{\lambda=0} = span\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

c) Therefore only two of the eigenvectors contribute to the spectral decomposition of A (and our data is really only 2-dimensional, even though the data vectors lie in $\mathbb{R}^5$. This is the "dimension reduction" of PCA.) Verify the spectral decomposition of A, using the outer product. (filled in after class)

$$A = 16 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & -1 \end{bmatrix} + 4 \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix} + 0 + 0$$

$$= 8 \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} + 2 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 0 & 0 & -8 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ -8 & 0 & 0 & 8 \end{bmatrix} \checkmark$$

~~d) Find and plot the coordinates in $\mathbb{R}^2$ of the original four data vectors, with respect to the the subspace spanned by the two dominating eigenvectors. (In the general case where the other eigenvalues were small but non-zero, these would be the coordinates of the projections of the original data vectors, onto the span of the eigenvectors with large eigenvalues.) Compare the geometry of this coordinate picture in $\mathbb{R}^2$ with the geometry of the original four points in $\mathbb{R}^5$!~~ not quite right - see next page.

The 4 attribute vectors corresponded to 5 individuals. Reading across the rows of the original list, the attribute vectors in $\mathbb{R}^4$ for each of the 5 individuals are

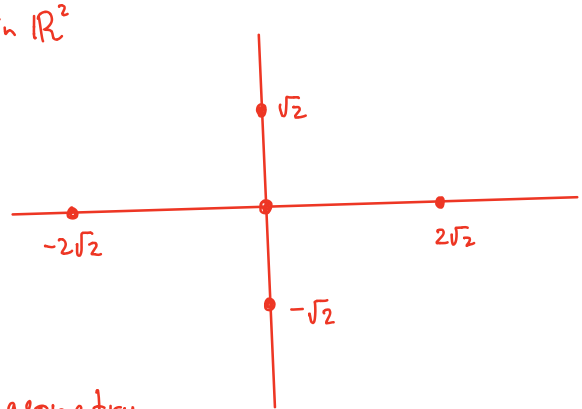
$$\left\{ \begin{bmatrix} -2 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

Find the coordinates of these vectors with respect to the orthonormal eigenbasis for the principal eigenvectors & sketch. How does this geometry compare to the geometry of the original 5 points in $\mathbb{R}^4$? (In the general case these would be the coordinates of the projection of the vectors onto the span of the principal eigenvectors).

<u>Ans</u> $\beta = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \right\}$

$$\begin{bmatrix} -2 \\ 0 \\ 0 \\ 2 \end{bmatrix}_\beta = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}_\beta = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}_\beta = \begin{bmatrix} 0 \\ -\sqrt{2} \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \end{bmatrix}_\beta = \begin{bmatrix} -2\sqrt{2} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}_\beta = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

<u>coord picture</u> in $\mathbb{R}^2$



in terms of geometry (distance & angle), these 5 points in the coordinate plane are congruent to the original 5 points, which were on a plane in $\mathbb{R}^4$. Geometrically, the original data set was 2-dim'l.