Today's topic: A **quadratic form** takes $\vec{\mathbf{x}}$ to $\langle \vec{\mathbf{x}}, A\vec{\mathbf{x}} \rangle = \vec{\mathbf{x}} \cdot A\vec{\mathbf{x}}$. Compare this to a linear transformation that takes $\vec{\mathbf{x}}$ to $A\vec{\mathbf{x}}$. Informally, linear transformations tell you coordinates, while quadratic forms tell you energy.

Important example: The sample mean M_n and the sample variance V_n are important statistics defined on a series of observations $\vec{\mathbf{x}}$. The sample mean is linear:

$$M_n = \tfrac{1}{n} \sum_{i=1}^{n} x_i \quad \text{(the average)}$$

The sample variance is a quadratic form:

$$V_n = \tfrac{1}{n-1} \sum_{i=1}^{n} (x_i - M_n)^2$$

For $n = 2$, we get

$$V_2(x, y) = \tfrac{1}{1} \left((x - (x+y)/2)^2 + (y - (x+y)/2)^2 \right) = \tfrac{1}{2} x^2 - xy + \tfrac{1}{2} y^2$$

This is the same as

$$V_2 = \vec{\mathbf{x}} \cdot A\vec{\mathbf{x}} \quad \text{for } \vec{\mathbf{x}} = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and } A = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

Is there a simpler formula for V_2? Yes, there is and we'll use chapter 5-7 to find it! The formula $\frac{1}{2}x^2 - xy + \frac{1}{2}y^2$ is ugly because "x" and "y" are the wrong variables. We need to choose better variables; we need to choose eigenvectors.

Find the eigenpairs of A. I'll call them $(c_0, \vec{\mathbf{v}}_0)$ and $(c_1, \vec{\mathbf{v}}_1)$.

Every $\vec{\mathbf{x}} = (x, y)$ can be written as the sum of two vectors: $\frac{\vec{\mathbf{v}}_0 \cdot \vec{\mathbf{x}}}{\vec{\mathbf{v}}_0 \cdot \vec{\mathbf{v}}_0} \vec{\mathbf{v}}_0$ added to $\frac{\vec{\mathbf{v}}_1 \cdot \vec{\mathbf{x}}}{\vec{\mathbf{v}}_1 \cdot \vec{\mathbf{v}}_1} \vec{\mathbf{v}}_1$. Simplify those two vectors using your values for $\vec{\mathbf{v}}_i$.

That means that

Simplify this using your values for a, b, c_i, and $\vec{\mathbf{v}}_i$.

$$\begin{aligned} V_2 = \vec{\mathbf{x}}^T A \vec{\mathbf{x}} &= (a\vec{\mathbf{v}}_0^T + b\vec{\mathbf{v}}_1^T)\left(A(a\vec{\mathbf{v}}_0 + b\vec{\mathbf{v}}_1)\right) \\ &= (a\vec{\mathbf{v}}_0^T + b\vec{\mathbf{v}}_1^T)(ac_0\vec{\mathbf{v}}_0 + bc_1\vec{\mathbf{v}}_1) \\ &= a^2 c_0 \|\vec{\mathbf{v}}_0\|^2 + b^2 c_1 \|\vec{\mathbf{v}}_1\|^2. \end{aligned}$$

Main Idea of 7.2: The **principal axes** of a quadratic form are the eigenvectors of its associated symmetric matrix. If $\vec{\mathbf{x}} \cdot A\vec{\mathbf{x}}$ is the quadratic form, $(c_i, \vec{\mathbf{v}}_i)$ are its eigenpairs, then the spectral decomposition of A is

$$A = \sum_{i=1}^{n} \frac{c_i}{\vec{\mathbf{v}}_i \cdot \vec{\mathbf{v}}_i} \vec{\mathbf{v}}_i \vec{\mathbf{v}}_i^T \quad \text{and} \quad \vec{\mathbf{x}} \cdot A\vec{\mathbf{x}} = \sum_{i=1}^{n} \frac{c_i}{\vec{\mathbf{v}}_i \cdot \vec{\mathbf{v}}_i} (\vec{\mathbf{x}} \cdot \vec{\mathbf{v}}_i)^2$$

Important example: Let's examine sample variance for n samples, instead of just 2.

$$V_n = \vec{\mathbf{x}} \cdot A\vec{\mathbf{x}} \quad \text{for } \vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and } A = \tfrac{1}{n-1} I - \tfrac{1}{n(n-1)} J$$

where I is the identity matrix and J is the matrix of all ones (a very famous matrix).

We can find an orthogonal basis of eigenvectors using chapter 5: The first eigenpair is $(c_1, \vec{\mathbf{v}}_1) = (0, \vec{\mathbf{j}})$ where $\vec{\mathbf{j}}$ is the vector of all ones. The other eigenvalues are all $\frac{1}{n-1}$, so we have to be a little clever to find orthogonal eigenvectors. I'll use the output of Gram-Schmidt when applied to the consuecutive difference eigenvectors $\vec{\mathbf{e}}_{i-1} - \vec{\mathbf{e}}_i$. We get the ith eigenpair is $(c_i, \vec{\mathbf{v}}_i)$ where $c_i = \frac{1}{n-1}$ and $\vec{\mathbf{v}}_i = \sum_{j=1}^{i-1} (e_j - e_i)$ is the vector that is all ones up to the $(i-1)$st coordinate, is $1 - i$ as the ith coordinate, and zero afterwards.

Here are the relevant dot products (here $i = 2, 3, \ldots, n$):

$$\begin{aligned} \vec{\mathbf{x}} \cdot \vec{\mathbf{v}}_1 &= x_1 + x_2 + \ldots + x_n && = nM_n \\ \vec{\mathbf{v}}_1 \cdot \vec{\mathbf{v}}_1 &= 1^2 + 1^2 + \ldots + 1^2 && = n \\ \vec{\mathbf{x}} \cdot \vec{\mathbf{v}}_i &= x_1 + x_2 + \ldots + x_{i-1} - (i-1)x_i && = (i-1)(M_{i-1} - x_i) \\ \vec{\mathbf{v}}_i \cdot \vec{\mathbf{v}}_i &= 1^2 + 1^2 + \ldots + 1^2 + (i-1)^2 = (i-1) + (i-1)^2 && = i(i-1) \end{aligned}$$

Using the principal axes of A we get

$$V_n = \vec{\mathbf{x}} \cdot A\vec{\mathbf{x}} = \sum_{i=1}^{n} \frac{c_i}{\vec{\mathbf{v}}_i \cdot \vec{\mathbf{v}}_i} (\vec{\mathbf{x}} \cdot \vec{\mathbf{v}}_i)^2 = \tfrac{1}{n-1} \sum_{i=2}^{n} (M_{i-1} - x_i)^2 (1 - 1/i)$$

Notice that

$$V_{n+1} = (1 - 1/n) V_n + (M_n - x_{n+1})^2 / (n+1)$$

This gives a simple, numerically stable, memory-less algorithm, published as Welford (1962), to update a running variance:

```
function [ new_count, new_ave, new_var] = init( new_obs )
    new_count = 1
    new_ave = new_obs
    new_var = 0
function [ new_count, new_ave, new_var] = update( old_count, old_ave, old_var, new_obs )
    new_count = old_count + 1;
    new_ave = old_ave + (new_obs - old_ave)/new_count
    new_var = old_var*(1-1/old_count) + ( old_ave - new_obs )^2/(new_count);
```

This demonstrates how eigenvectors can be unexpectedly useful (especially in real time computing such as aircraft control).

MA322-001 Apr 21 Quiz Name: ____________________

7.1: Let A be the diagonal matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

(a) Compute the eigenpairs $(c_i, \vec{\mathbf{v}}_i)$ of A:

(b) Compute the rank one matrices $\frac{c_i}{\vec{\mathbf{v}}_i \cdot \vec{\mathbf{v}}_i} \vec{\mathbf{v}}_i \vec{\mathbf{v}}_i^T$

(c) Add them up to get the spectral decomposition of A:

7.2: Find a matrix B so that if $\vec{\mathbf{x}} = \begin{bmatrix} x \\ y \end{bmatrix}$ then $\vec{\mathbf{x}} \cdot B\vec{\mathbf{x}} = x^2 + 14xy + 5y^2$.

(b) Actually multiply out $B\vec{\mathbf{x}}$

(c) Actually multiply out $\vec{\mathbf{x}} \cdot B\vec{\mathbf{x}}$