

Key

Today's topic: A quadratic form takes \vec{x} to $\langle \vec{x}, A\vec{x} \rangle = \vec{x} \cdot A\vec{x}$. Compare this to a linear transformation that takes \vec{x} to $A\vec{x}$. 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{x} . The sample mean is linear:

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

The sample variance is a quadratic form:

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

For $n = 2$, we get

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

This is the same as

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

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{v}_0) and (c_1, \vec{v}_1) .

$$\left(0, \begin{bmatrix} x \\ x \end{bmatrix} \right), \left(1, \begin{bmatrix} x \\ -x \end{bmatrix} \right) \quad x \neq 0$$

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

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{x+y}{2} \\ \frac{x+y}{2} \end{bmatrix} + \begin{bmatrix} \frac{x-y}{2} \\ \frac{y-x}{2} \end{bmatrix}$$

$\vec{x}_s \quad \quad \quad \vec{x}_a$

That means that

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

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

$$\begin{aligned} (\vec{x}_s + \vec{x}_a) \cdot A(\vec{x}_s + \vec{x}_a) &= \vec{x}_s \cdot A\vec{x}_s + \vec{x}_a \cdot A\vec{x}_a \\ &= 0 + 2\left(\frac{x-y}{2}\right)^2 \\ &= \frac{(x-y)^2}{2} \end{aligned}$$

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

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

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

$$V_n = \vec{x} \cdot A \vec{x} \quad \text{for } \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad A = \frac{1}{n-1} I - \frac{1}{n(n-1)} J = \begin{bmatrix} \frac{1}{n} & -\frac{1}{n(n-1)} & \dots & 0 \\ -\frac{1}{n(n-1)} & \frac{1}{n} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n(n-1)} & \dots & \dots & \frac{1}{n} \end{bmatrix}$$

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 $(c_1, \vec{v}_1) = (0, \vec{j})$ where \vec{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 consecutive difference eigenvectors $\vec{e}_{i-1} - \vec{e}_i$. We get the i th eigenpair is (c_i, \vec{v}_i) where $c_i = \frac{1}{n-1}$ and $\vec{v}_i = \sum_{j=1}^{i-1} (e_j - e_i)$ is the vector that is all ones up to the $(i-1)$ st coordinate, is $1-i$ as the i th coordinate, and zero afterwards. $\vec{v}_5^T = [1, 1, 1, 1, -4, 0, 0, 0, \dots]$

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

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

Using the principal axes of A we get

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

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).

Name: Key

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{v}_i) of A :

$$\left(1, \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}\right), \left(3, \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix}\right), \left(5, \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}\right)$$

x, y, z
any non zero
numbers (like $x=y=z=1$)

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

$$\frac{1}{x^2 + 0 + 0} \begin{bmatrix} x \cdot x & x \cdot 0 & x \cdot 0 \\ 0 \cdot x & 0 \cdot 0 & 0 \cdot 0 \\ 0 \cdot x & 0 \cdot 0 & 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \frac{3}{0 + y^2 + 0} \begin{bmatrix} 0 & 0 & 0 \\ 0 & y^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

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

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} = A$$

and $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

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

$$B = \begin{bmatrix} 1 & 7 \\ 7 & 5 \end{bmatrix}$$

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

$$B\vec{x} = \begin{bmatrix} 1 & 7 \\ 7 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 7y \\ 7x + 5y \end{bmatrix}$$

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

$$\begin{aligned} \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} x + 7y \\ 7x + 5y \end{bmatrix} &= \begin{bmatrix} x(x + 7y) + y(7x + 5y) \end{bmatrix} \\ &= \begin{bmatrix} x^2 + 7xy + 7yx + 5y^2 \end{bmatrix} \\ &= \begin{bmatrix} x^2 + 14xy + 5y^2 \end{bmatrix} \end{aligned}$$

