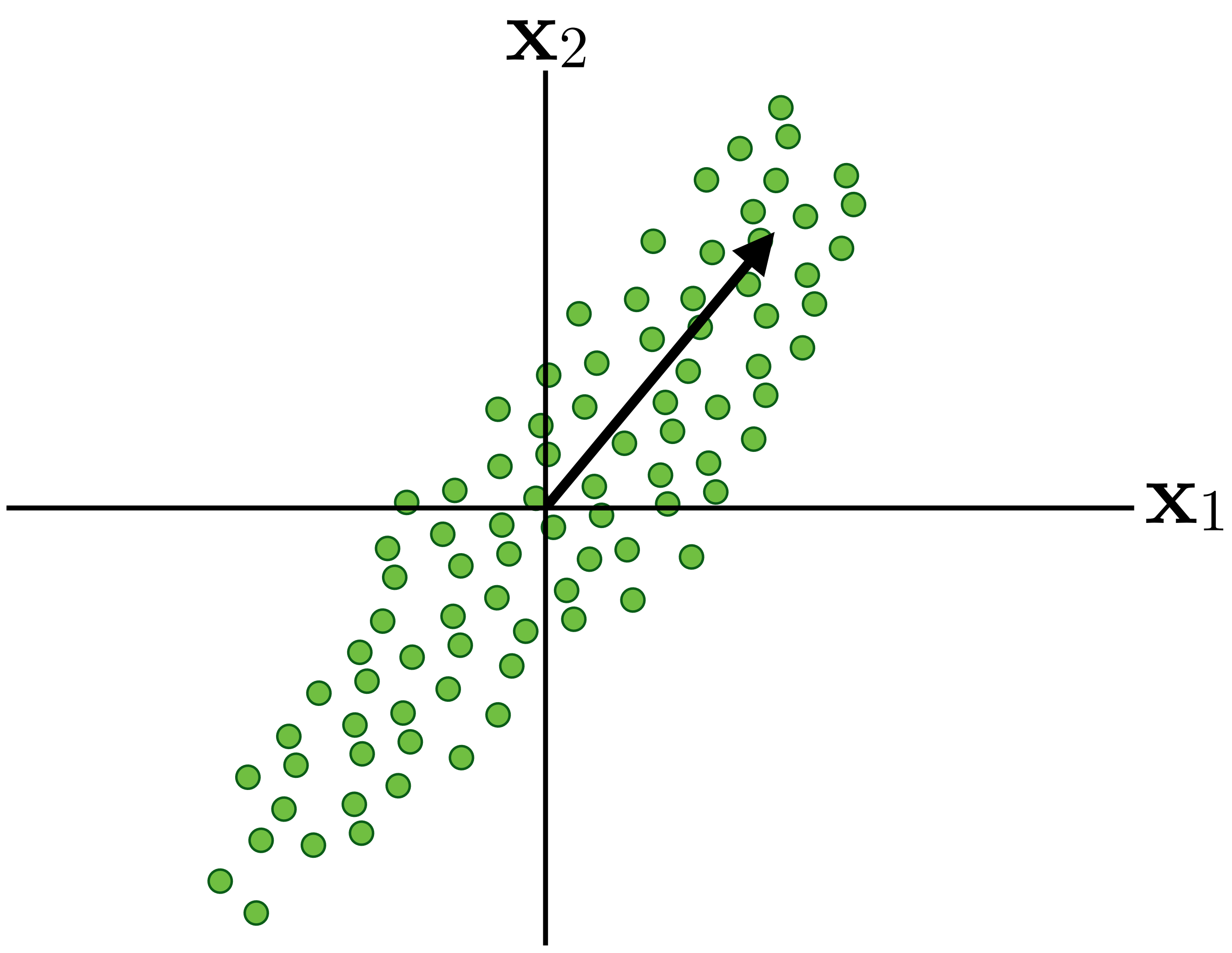
Eigenvalues, Eigenvectors, and an Intro to PCA

Changing Basis

- Talked about re-writing our data using a new set of variables, or a new basis
- How do we choose this new basis?

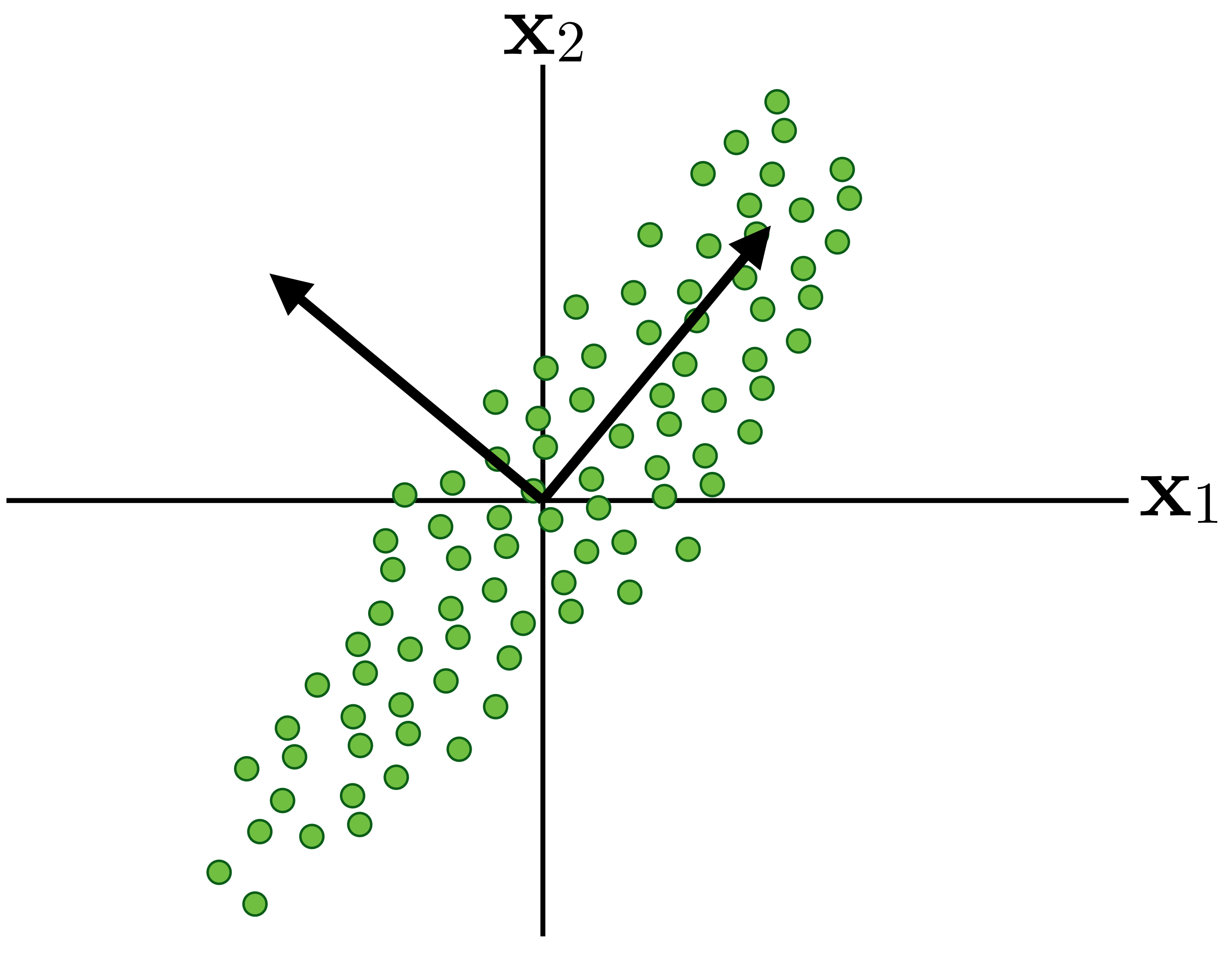
PCA

One (very popular) method: start by choosing the basis vectors as directions in which the variance of the data is maximal.



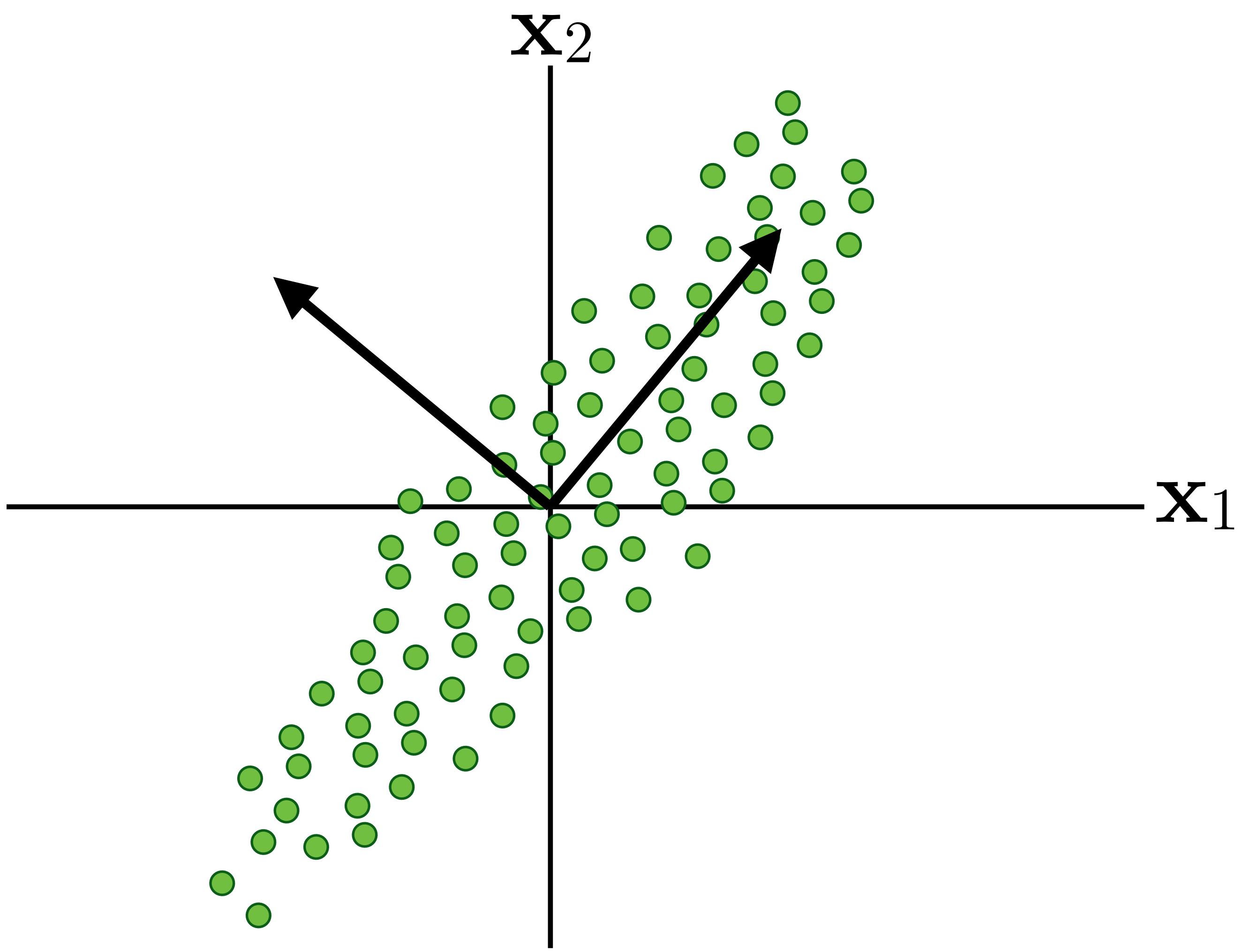
PCA

Then, choose subsequent directions that are orthogonal to the first and have *next* largest variance.



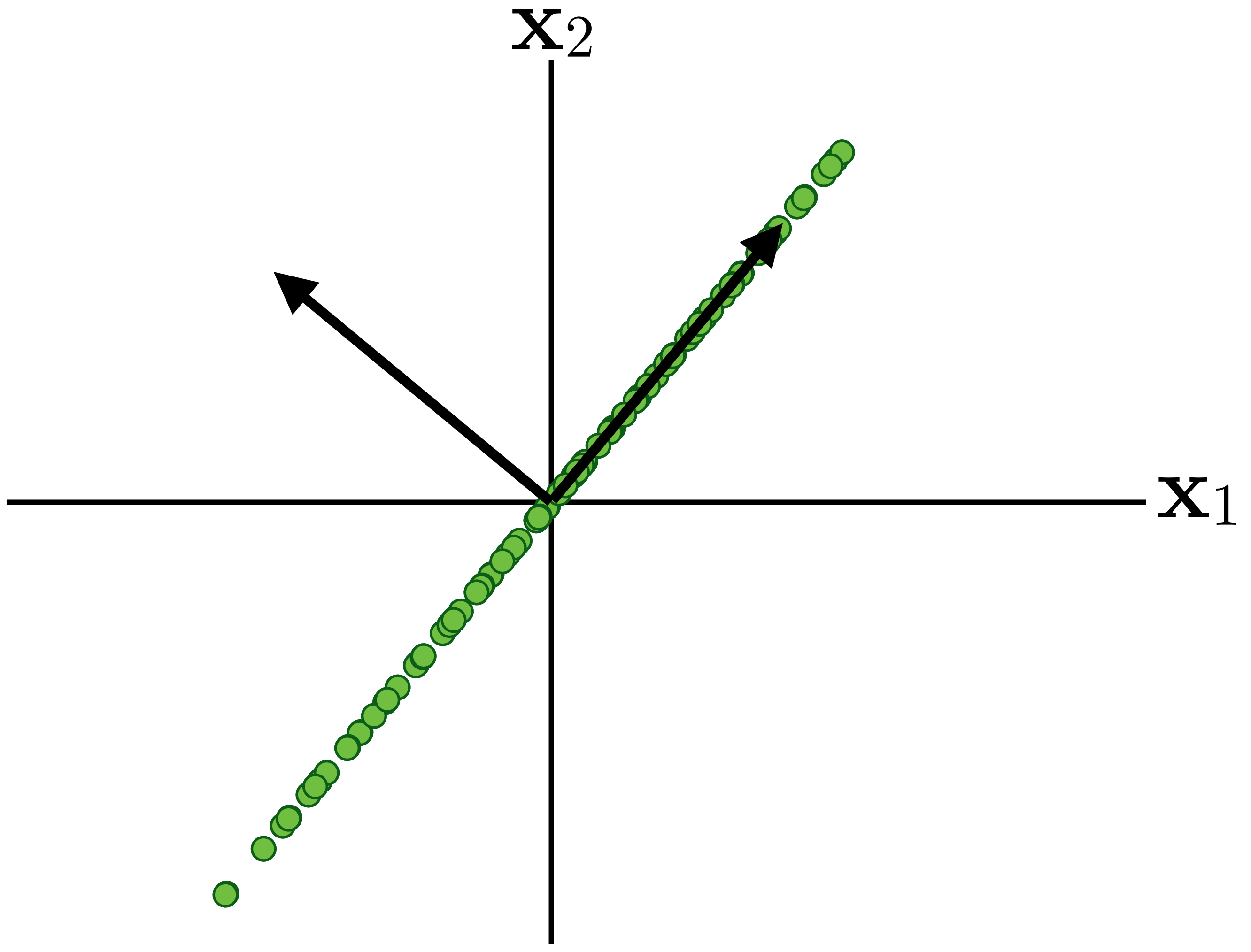
PCA

The variance in a given direction refers to the variance of the data once projected onto that direction.



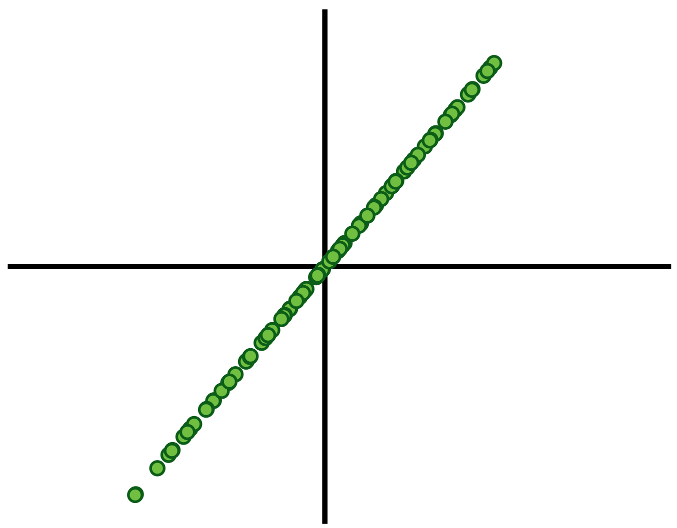
Directional Variance

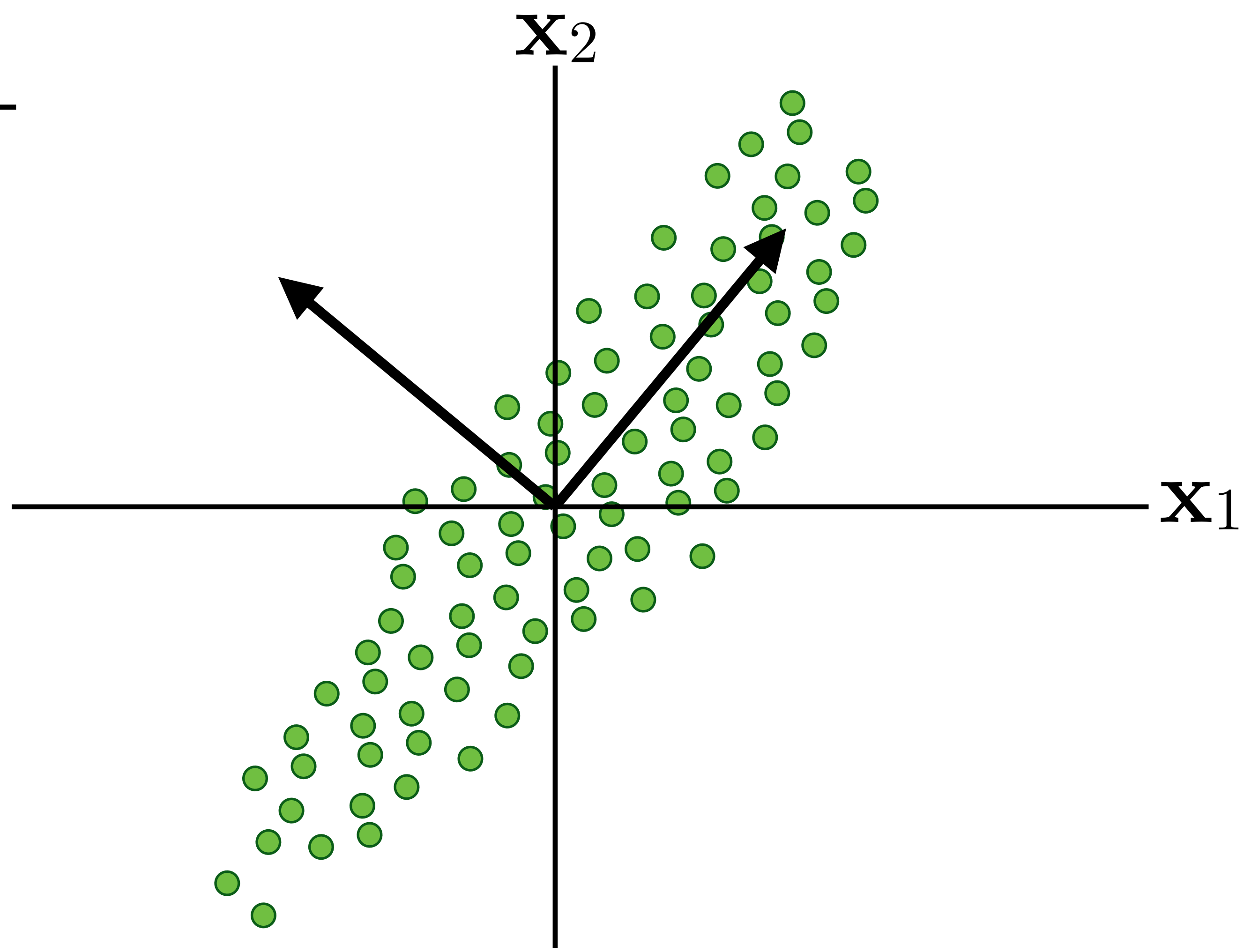
The variance in a given direction refers to the variance of the data once projected onto that direction.



Directional Variance

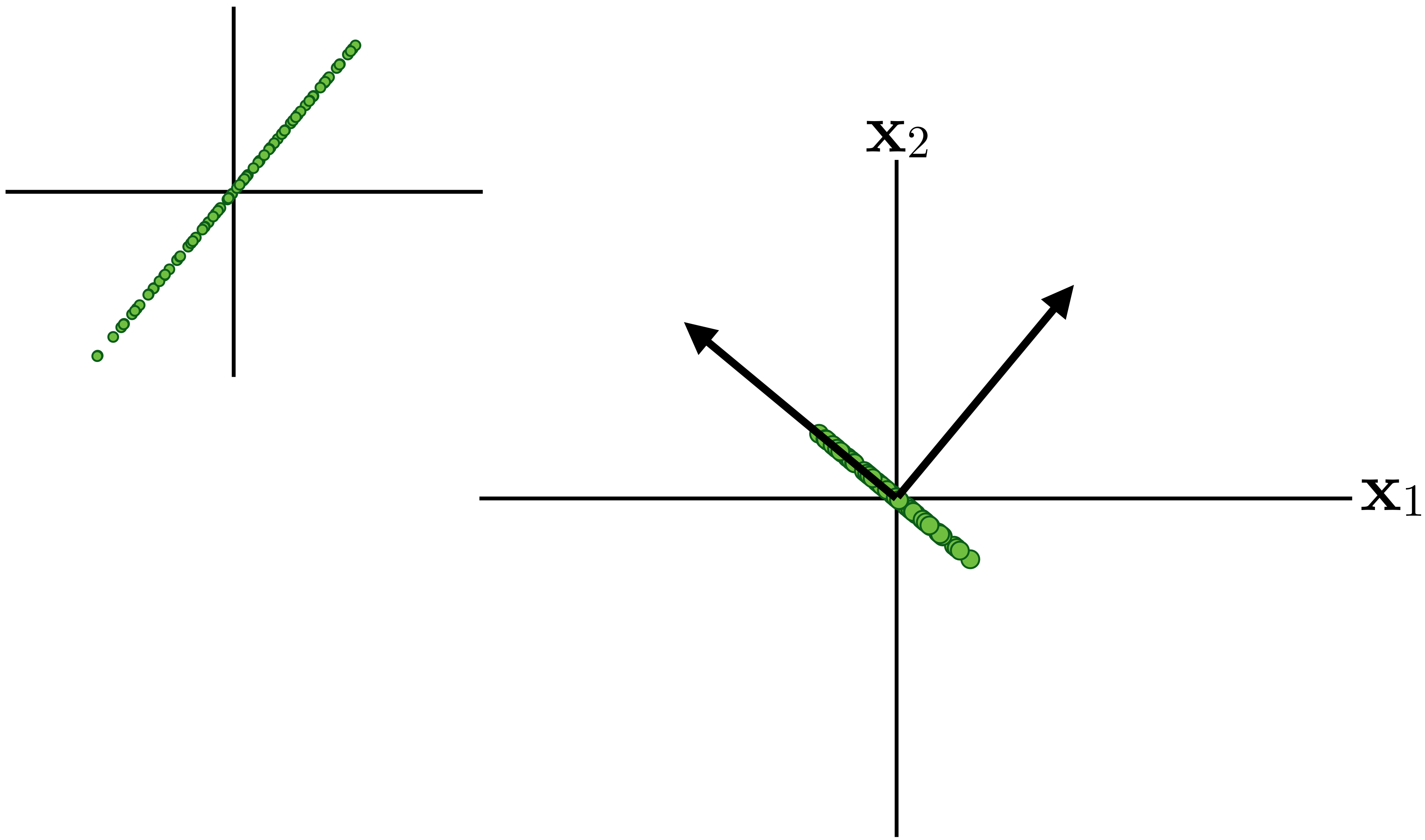
The variance in a given direction refers to the variance of the data once projected onto that direction.





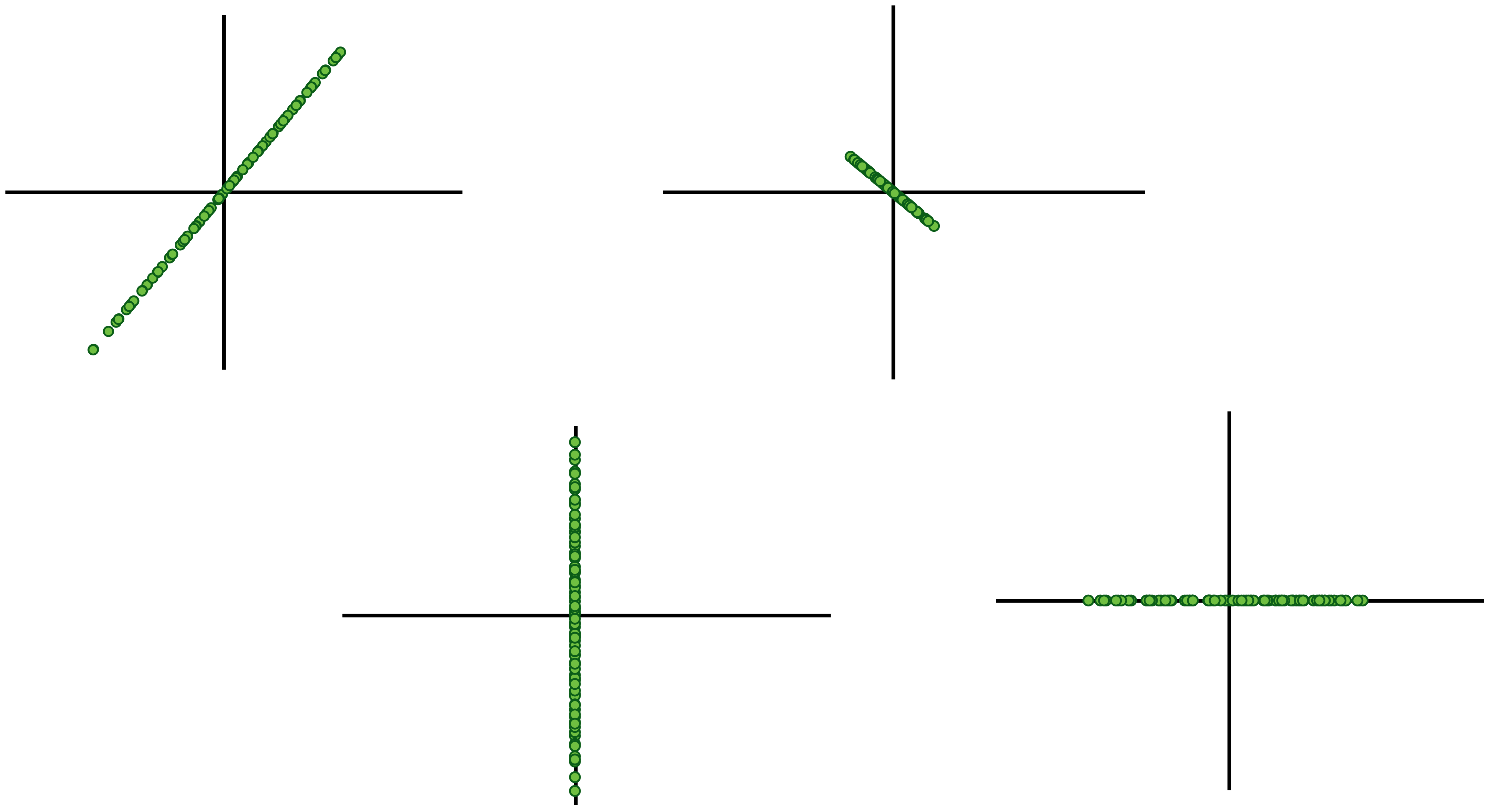
Directional Variance

The variance in a given direction refers to the variance of the data once projected onto that direction.



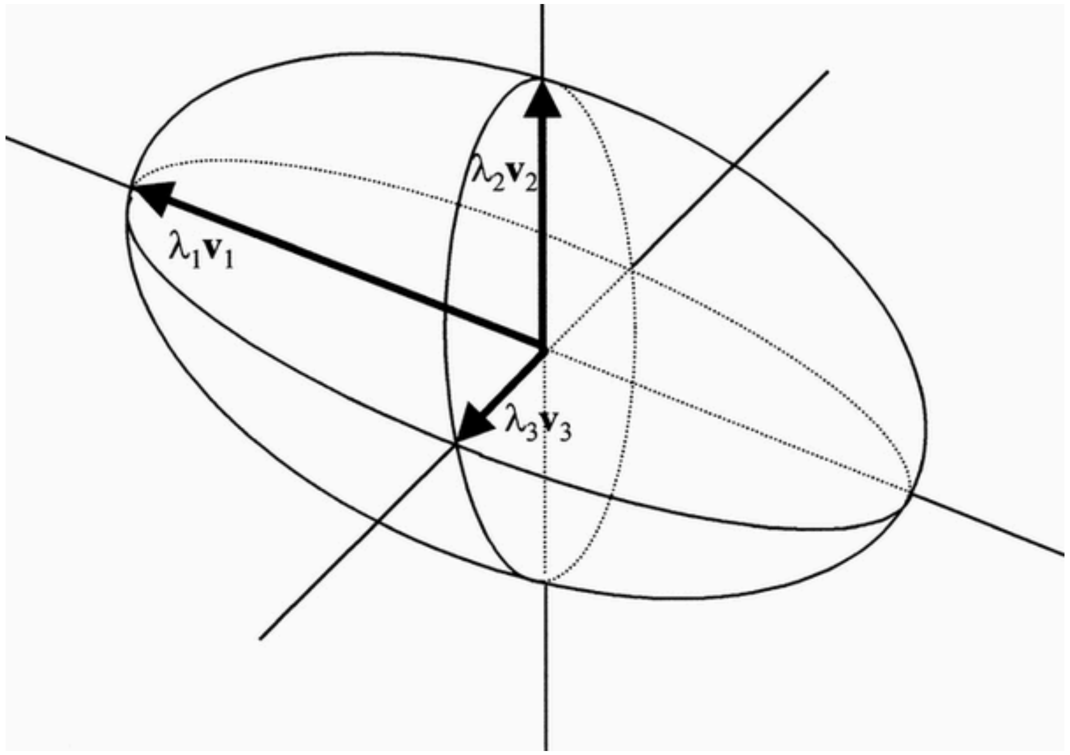
Directional Variance

The variance in a given direction refers to the variance of the data once projected onto that direction.



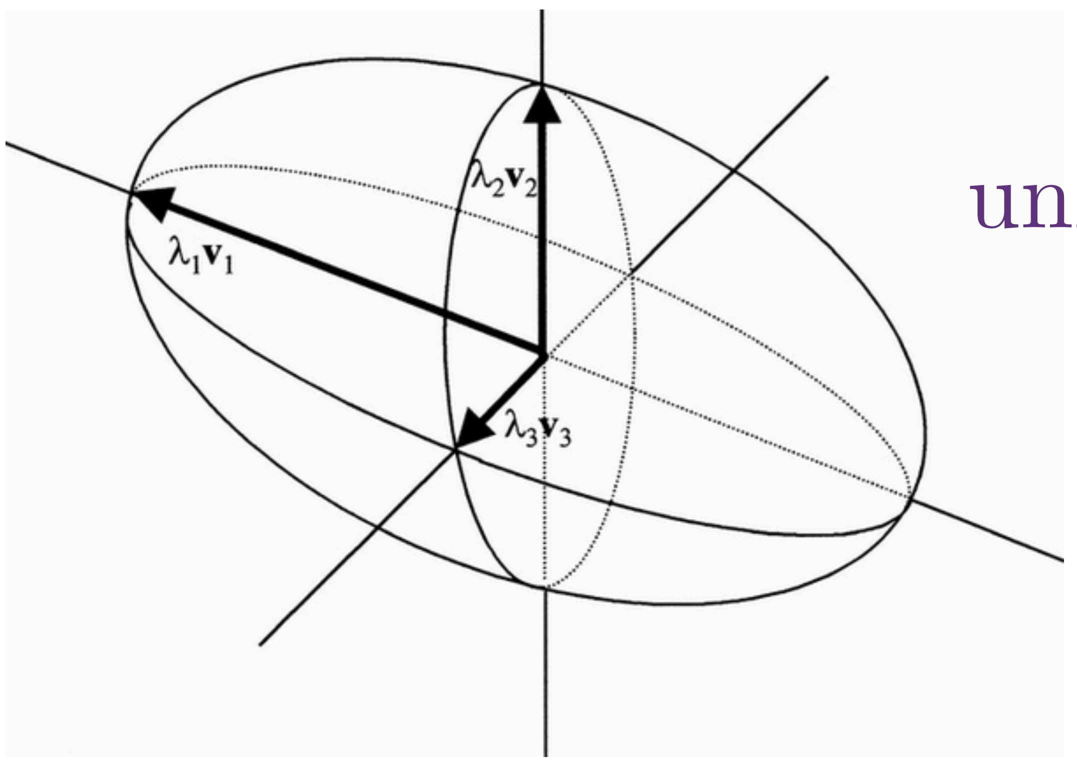
Directional Variance

Eigenvectors of the covariance matrix provide these directions of maximal variance.



They are the major/minor axes of the ellipsoid associated with the elliptical distribution

Directional Variance



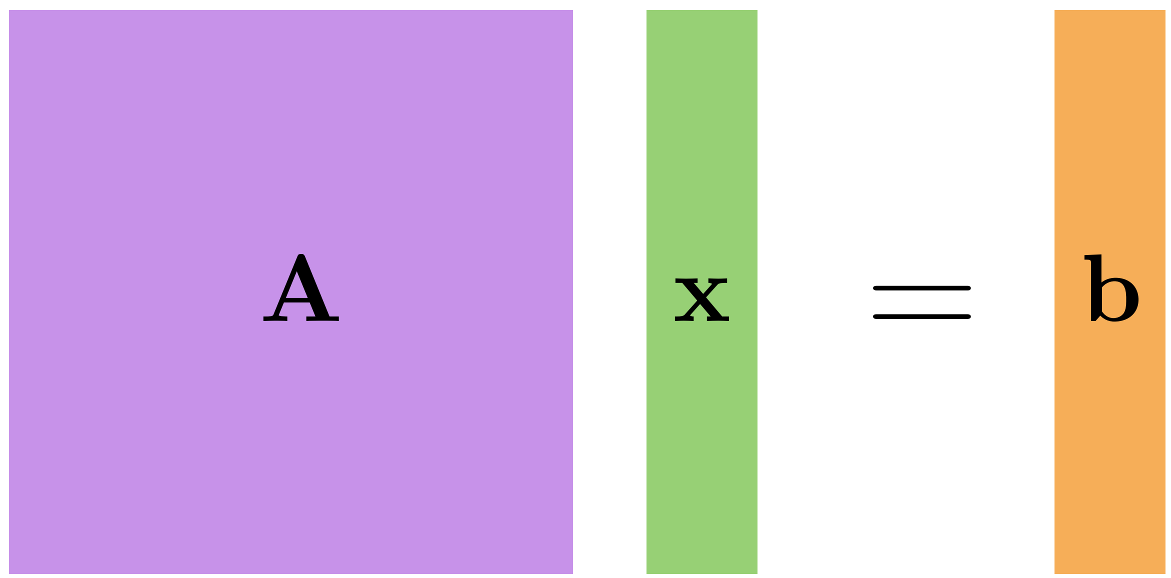
The $\mathbf{v_i}$ vectors are
unit vectors specifying direction.

The λ_i are scalars indicating magnitude of
spread in each direction of each elliptical axis.

Eigenvectors

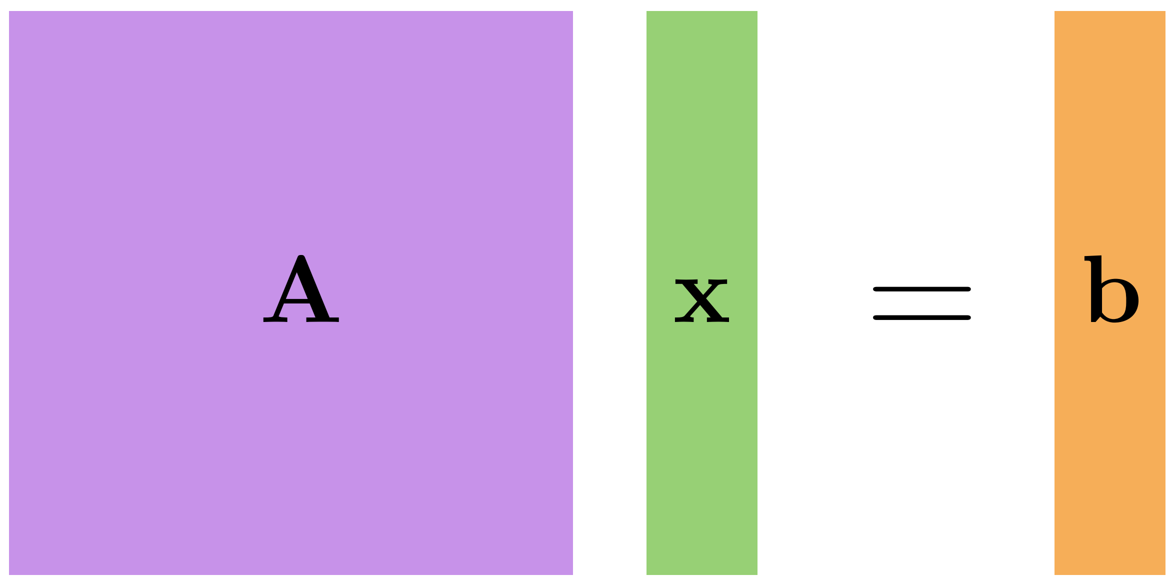
Effect of Matrix Multiplication

When we multiply a vector by a matrix, what do we get?



Another vector.

Effect of Matrix Multiplication



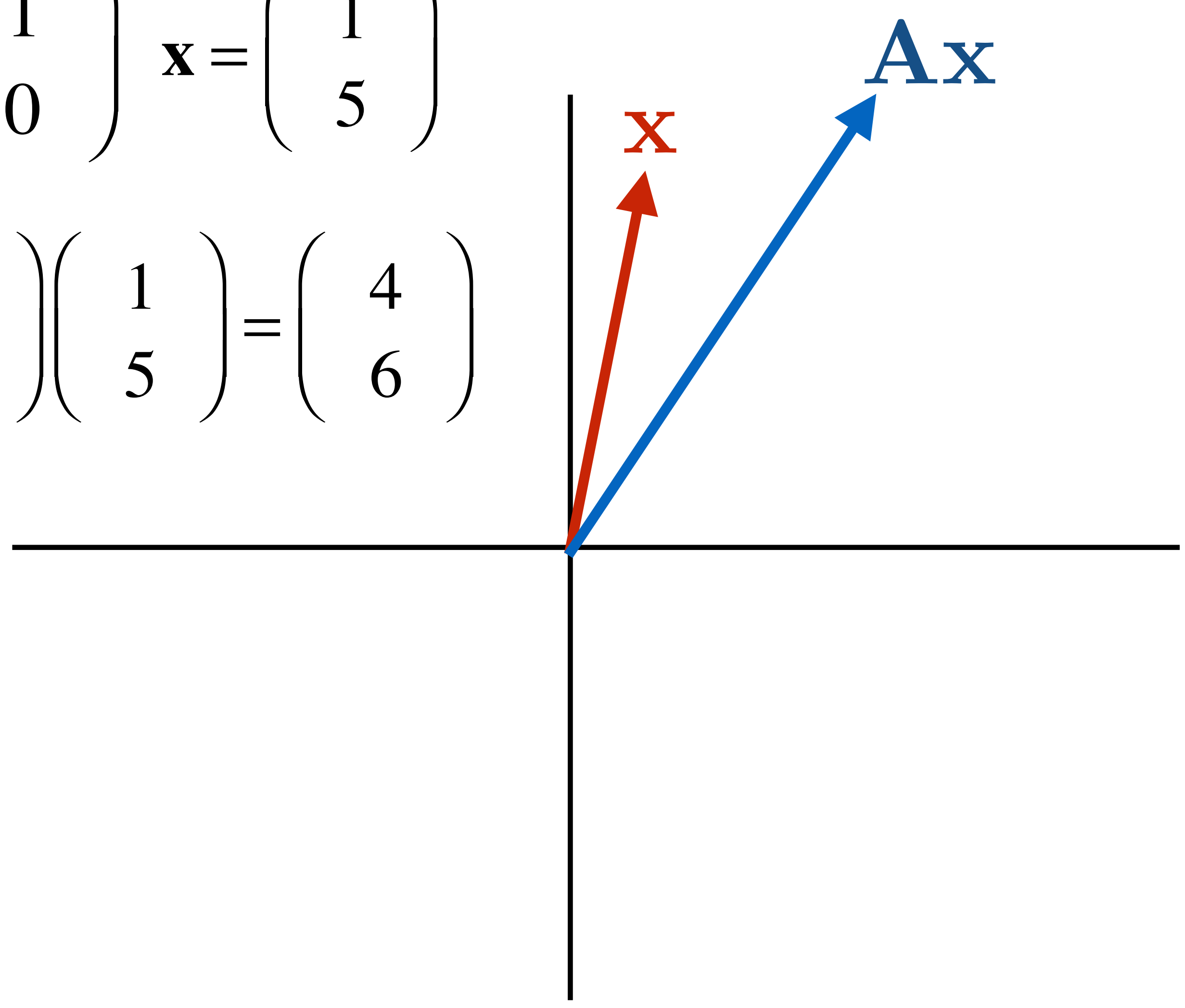
When the matrix $\mathbf{A}$ is square, then $\mathbf{x}$ and $\mathbf{b}$ have the same size. We can draw (or imagine) them in the same space.

Linear Transformation

In general, multiplying a vector by a matrix changes both its direction and magnitude.

$$\mathbf{A} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$$

$$\mathbf{Ax} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

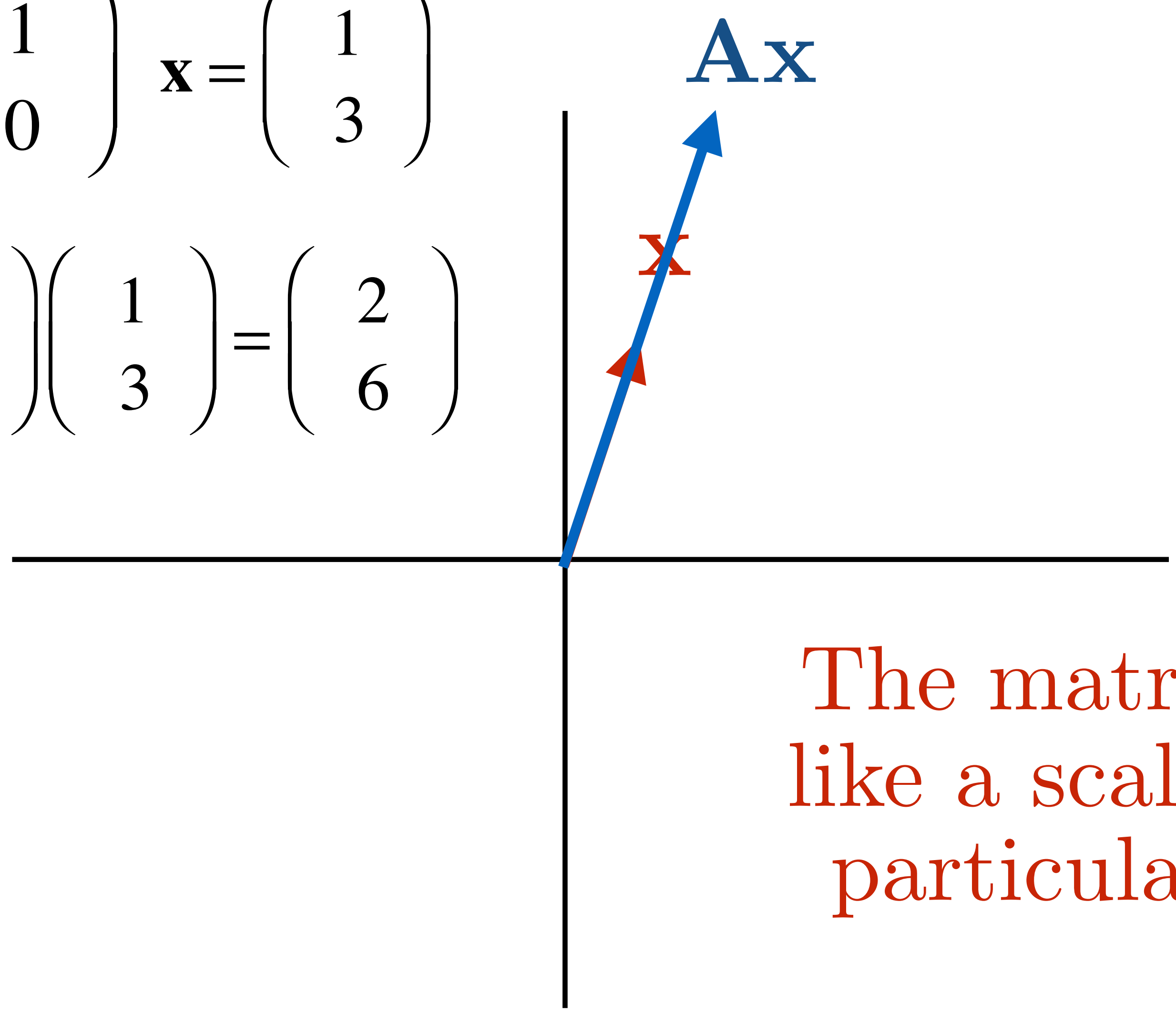


Eigenvectors

However, a matrix may act on certain vectors by changing *only* their magnitude, *not* their spanning direction.

$$\mathbf{A} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

$$\mathbf{Ax} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \end{pmatrix}$$



The matrix **A** acts like a scalar on this particular vector!

Eigenvalues and Eigenvectors

- For a square matrix $\mathbf{A}$, a nonzero vector $\mathbf{x}$ is called an **eigenvector of A** if multiplying by $\mathbf{A}$ results in a scalar multiple of $\mathbf{x}$.

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

- The scalar λ is called the **eigenvalue** associated with the eigenvector.

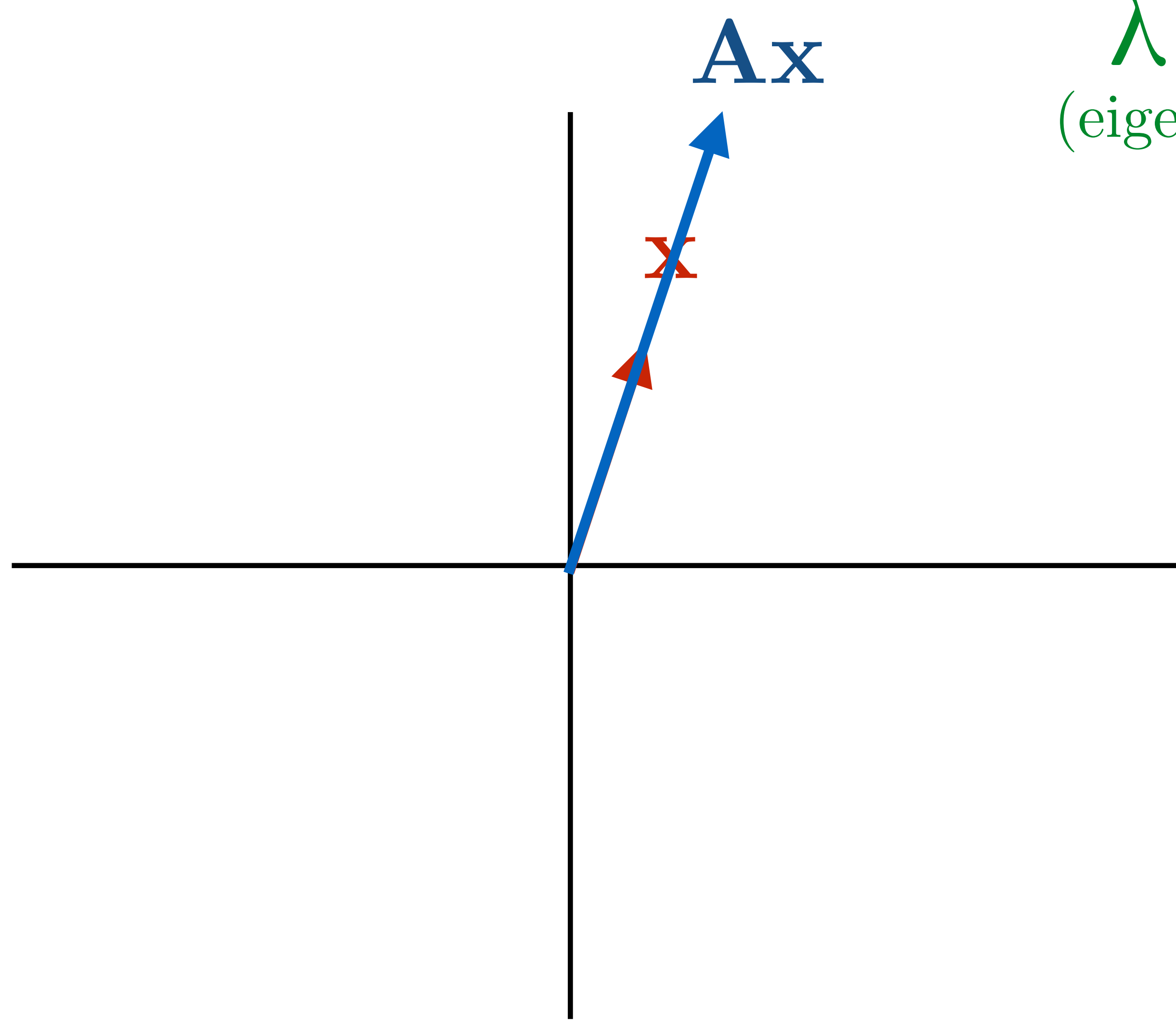
Previous Example

$$\mathbf{A} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

$$\mathbf{Ax} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 3 \end{pmatrix} = 2\mathbf{x}$$

$\lambda = 2$

(eigenvalue)

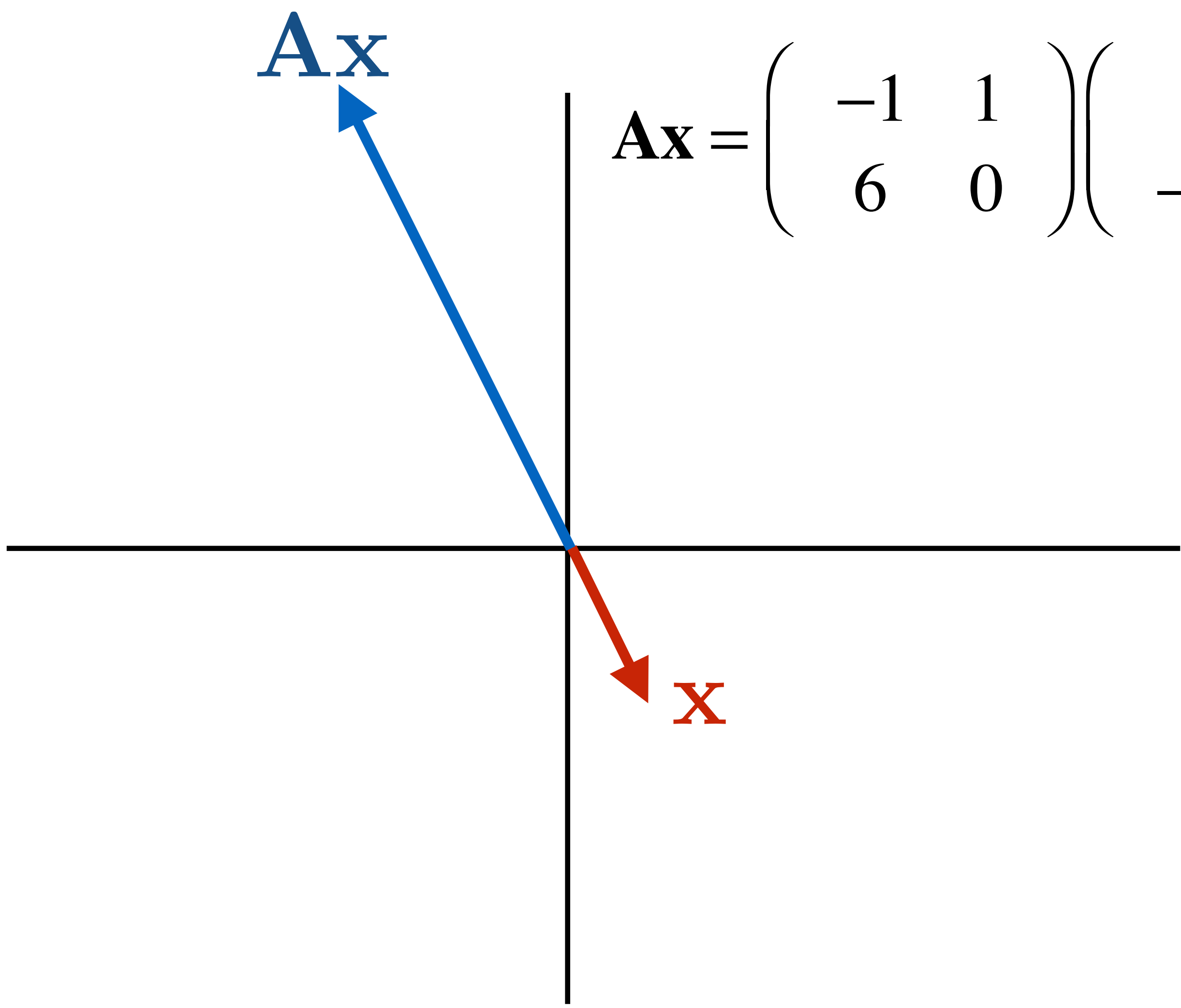


Example 2

Show that x is an eigenvector of A and find the corresponding eigenvalue.

$$\mathbf{A} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

$$\mathbf{Ax} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \end{pmatrix} =$$



Practice

1

Show that $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ and find the corresponding eigenvalue.

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 3 \\ -3 \end{pmatrix}$$

2

Can a rectangular matrix have eigenvalues/eigenvectors?

Eigenvector/Eigenvalue Facts

1. Only square matrices have eigenvectors.

2. Eigenvectors and eigenvalues come in pairs.

3. An nxn matrix has n **eigenpairs**, although some eigenvalues may be zero if the matrix is not full rank.

4. All square matrices have eigenvectors, but most of them will contain complex numbers ($i = \sqrt{-1}$)

5. The eigenvalues of a matrix are commonly called the **spectrum** of the matrix.

Eigenvector/Eigenvalue Facts

6. Any scalar multiple of an eigenvector of **A** is also an eigenvector of **A** with the same eigenvalue.

$$\mathbf{A} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \quad \lambda = 2$$

Try:

$$\mathbf{v} = \begin{pmatrix} 2 \\ 6 \end{pmatrix} \text{ or } \mathbf{u} = \begin{pmatrix} -1 \\ -3 \end{pmatrix} \text{ or } \mathbf{z} = \begin{pmatrix} 3 \\ 9 \end{pmatrix}$$

In general (#proof): Let $\mathbf{Ax} = \lambda\mathbf{x}$. If c is some constant, then:

$$\mathbf{A}(c\mathbf{x}) =$$

$c\mathbf{x}$ is also an eigenvector

Eigenspaces

- For a given matrix, there are infinitely many eigenvectors associated with one eigenvalue.
- Any scalar multiple (positive or negative) can be used.
- The collection is called the eigenspace associated with the eigenvalue.
- In previous example, the eigenspace associated with $\lambda = 2$ is span $\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}$
- With this in mind, what should you expect from software??

Zero Eigenvalues

What if $\lambda = 0$ is an eigenvalue for some matrix $\mathbf{A}$?

$$\mathbf{A}\mathbf{x} = 0\mathbf{x} = \mathbf{0}, \quad \text{where} \quad \mathbf{x} \neq \mathbf{0} \quad \text{is an eigenvector}$$

This means some linear combination of the columns of $\mathbf{A}$ is equal to zero!

$\Longrightarrow$ Columns of $\mathbf{A}$ are linearly dependent

$\Longrightarrow$ $\mathbf{A}$ is not full rank

$\Longrightarrow$ Perfect Multicollinearity

Eigenvalue Ordering

The eigenpairs (λ_i,$\mathbf{v}_i$) of a matrix are ordered by the *magnitude* of the eigenvalue.

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

The *"first"* eigenvector $\mathbf{v}_1$ is the eigenvector associated with the largest eigenvalue (in absolute value)

Practice

1

For the following matrix, determine the eigenvalue associated with the given eigenvector.

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & -1 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

2

What can you conclude about the matrix $\mathbf{A}$ from this?

3

The matrix $\mathbf{M}$ has eigenvectors $\mathbf{u}$ and $\mathbf{v}$ as shown. What is λ_1, the first eigenvalue of $\mathbf{M}$?

$$\mathbf{M} = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix} \mathbf{u} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \mathbf{v} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

Symmetric Matrices

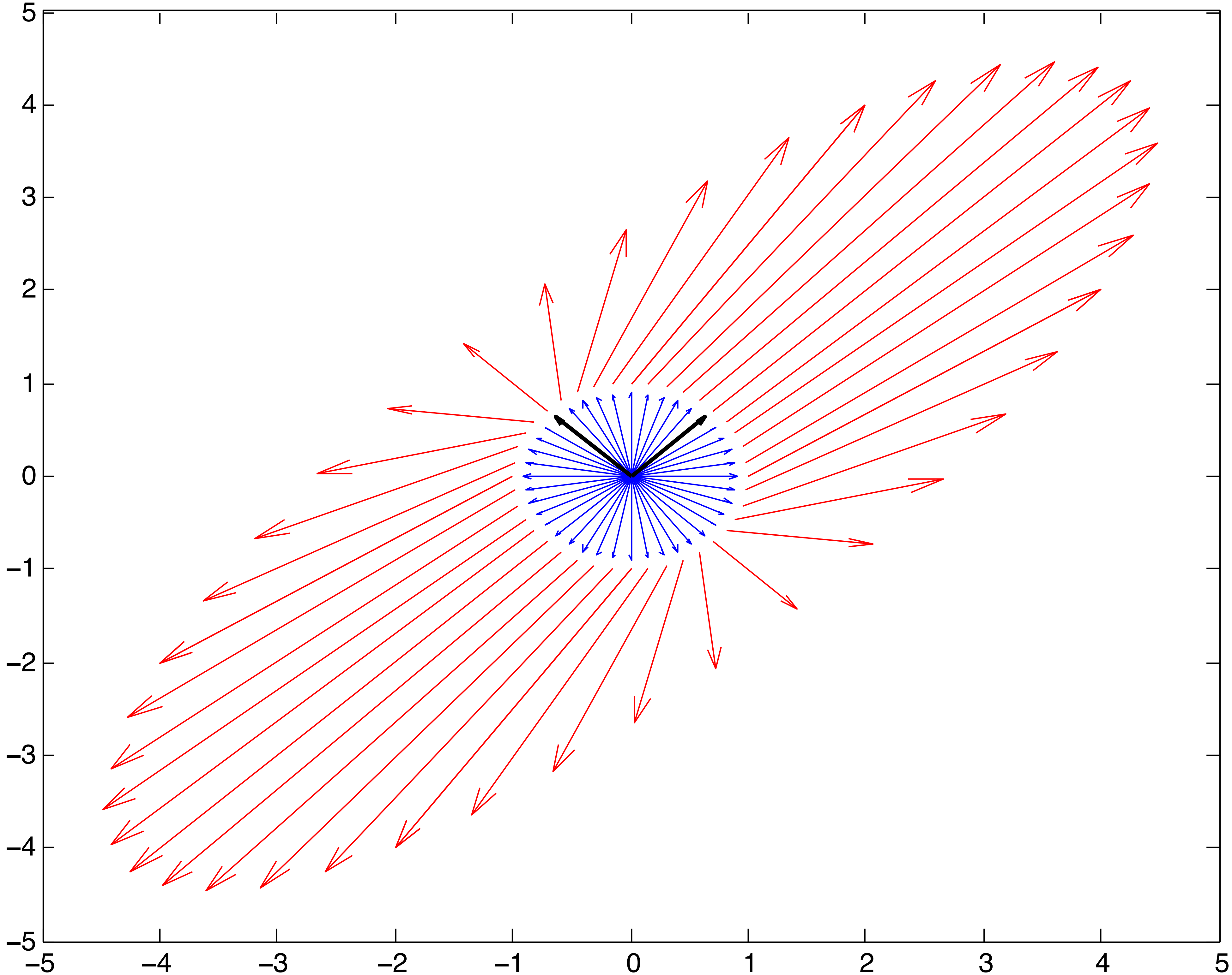
- Symmetric matrices (like the covariance, correlation, distance and similarity matrices) have several nice properties:

1. Their eigenvalues/eigenvectors are real (as opposed to complex ($i = \sqrt{-1}$))

2. Their eigenvectors are all mutually orthogonal.

3. Thus if you normalize the eigenvectors to unit length they will form an orthogonal matrix.

Eigenvectors of Symmetric Matrices



Introduction to Principal Components Analysis (PCA)

Eigenvectors of the Covariance/Correlation Matrices

- Covariance/Correlation Matrices are symmetric
- Their eigenvectors are orthogonal
- Eigenvectors are ordered by the magnitude of their eigenvalues
- Eigenvectors are assumed to be unit vectors, expressing only a direction.

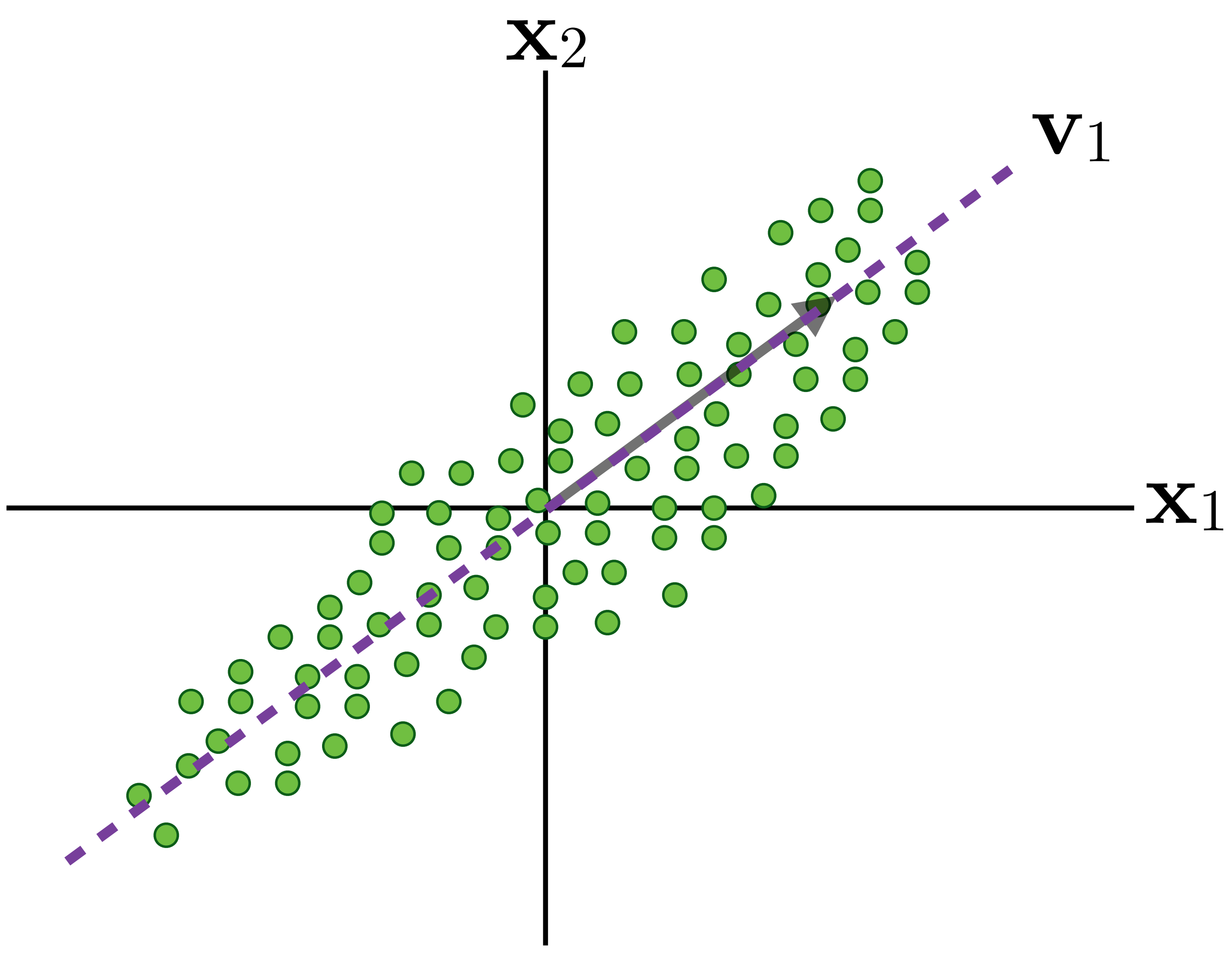
Covariance vs. Correlation

More detail later, but

- For the covariance matrix, we want to think of our data as *centered* to begin with (directions drawn from the origin=mean).
- For the correlation matrix, we want to think of our data as *standardized* to begin with (i.e. centered *and* divided by standard deviation)

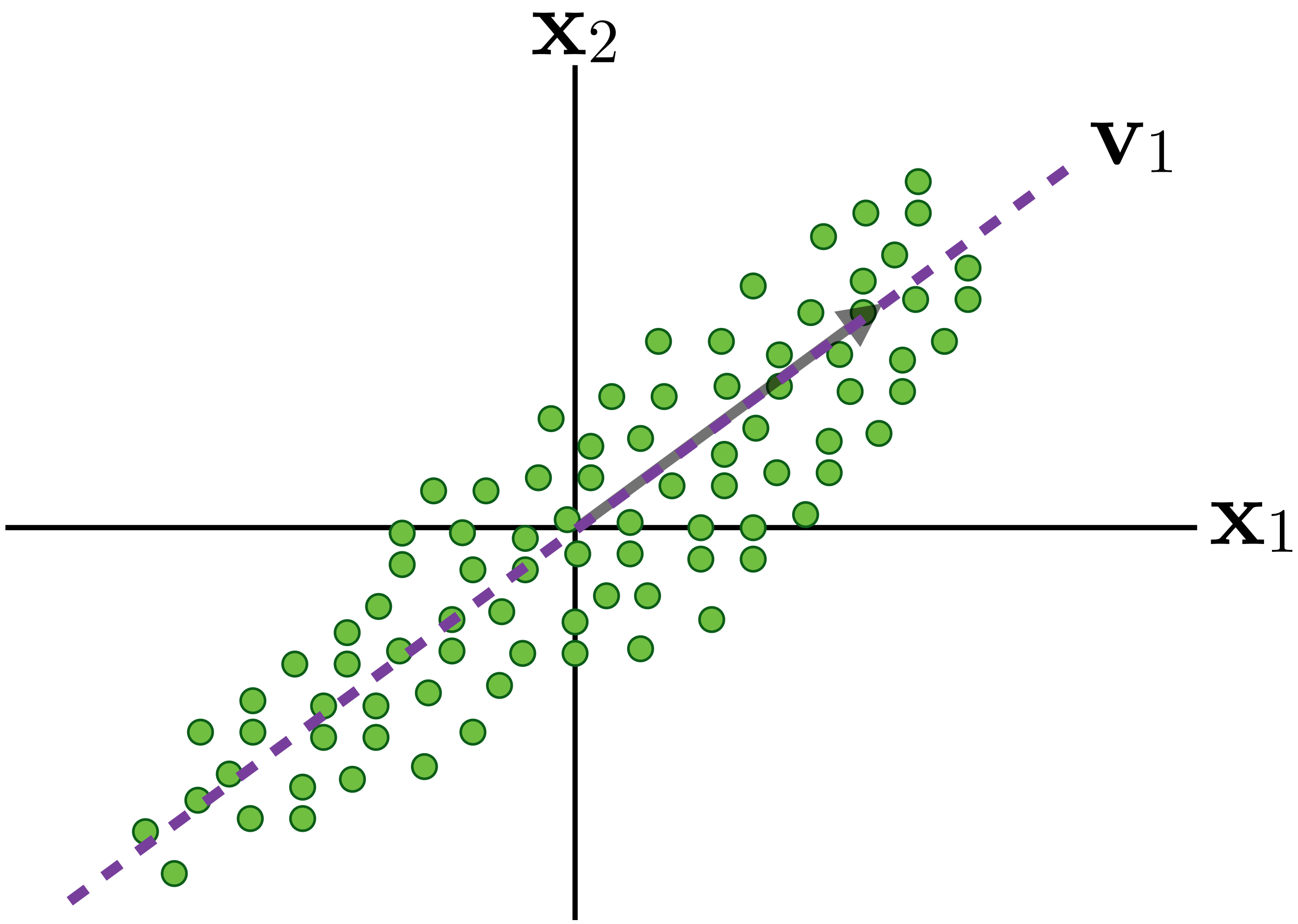
Direction of Maximal Variance

The first eigenvector of a covariance/correlation matrix points in the direction of maximum variance in the data. This eigenvector is the **first principal component**.



"Best" Approximation

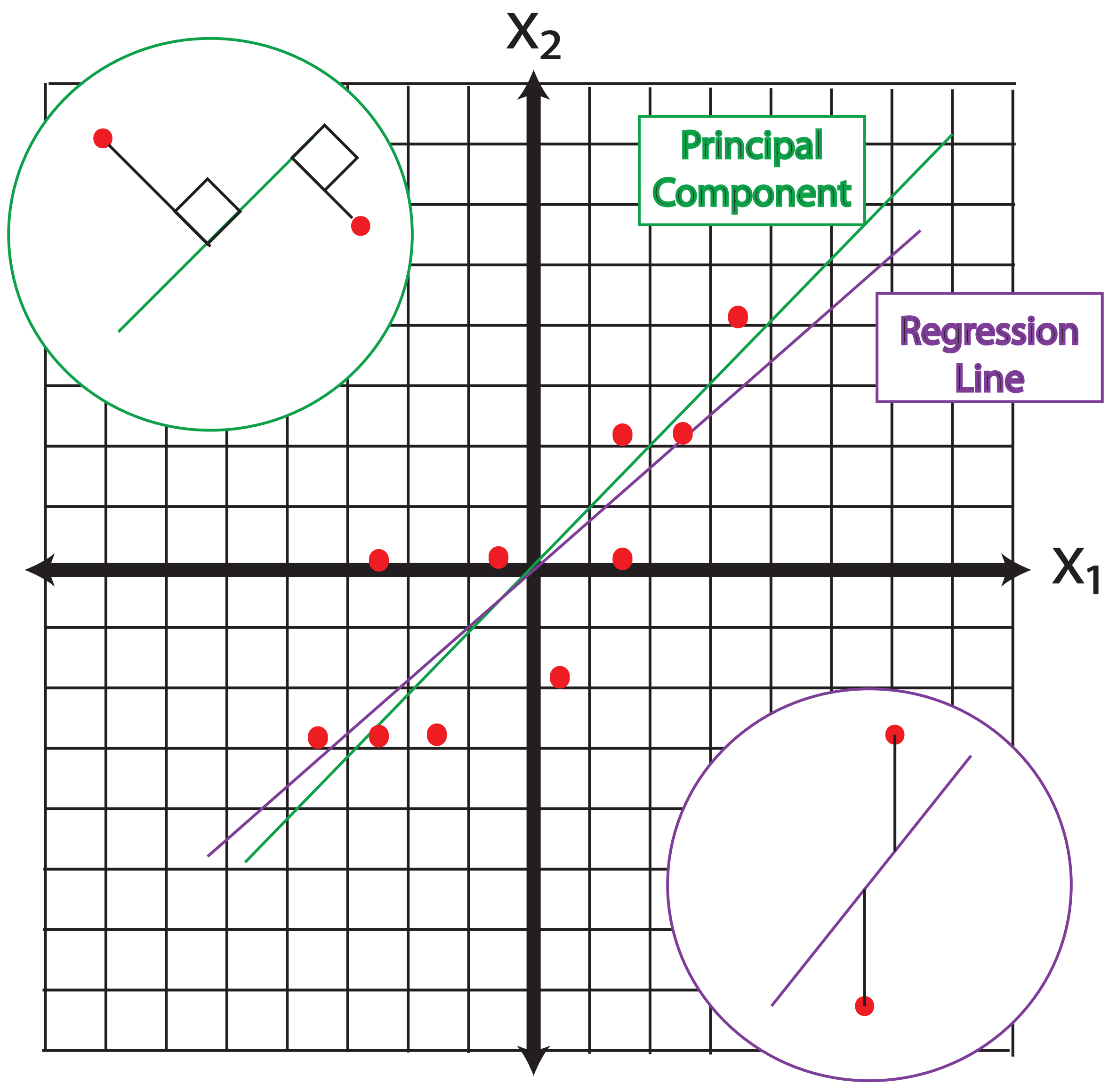
The first principal component minimizes the orthogonal distances between the spanning line and the points.



Projecting data onto this direction gives the best 1-dimensional approximation of the 2-dimensional data.

Not a regression line!

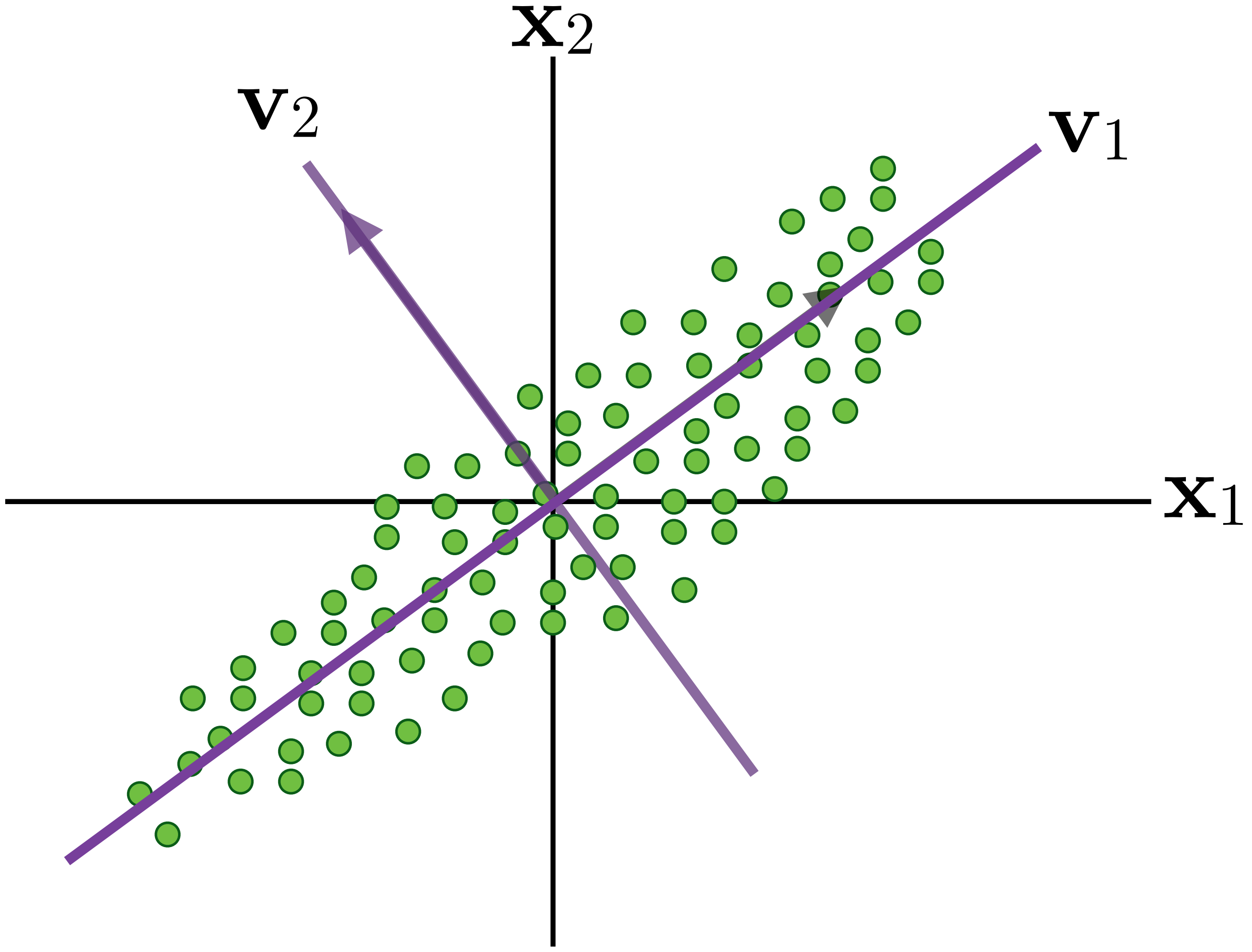
While it may look close in many two dimensional situations, there is no target variable in PCA.



Orthogonal Distances vs. Vertical Distances!

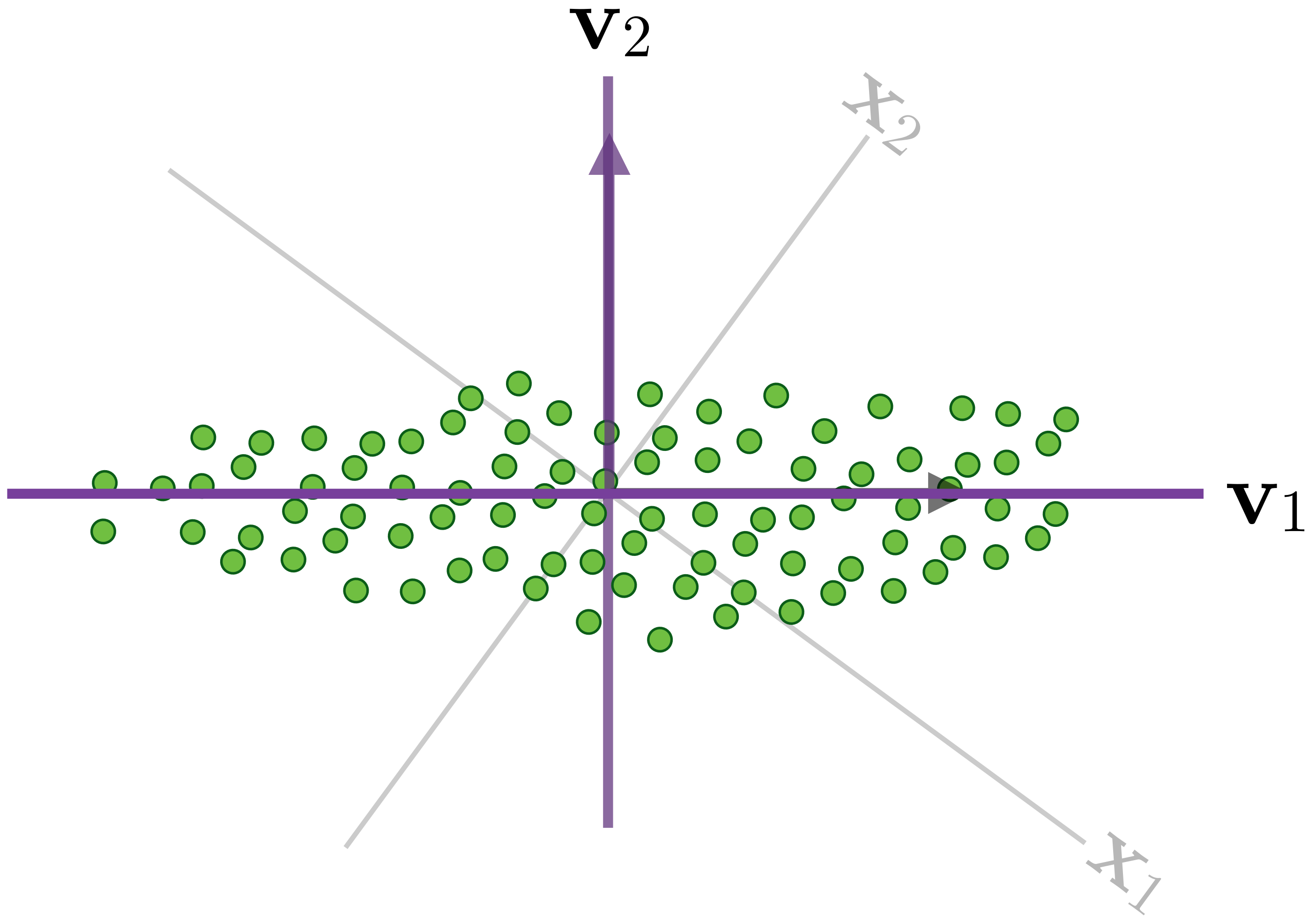
Secondary Directions

The second eigenvector of a covariance matrix points in the direction, orthogonal to the first, of maximal variance



A New Basis

Principal components provide us with a new orthogonal basis where the coordinates of the data points are uncorrelated.



Variable loadings

- Each principal component is a linear combination of the original variables:

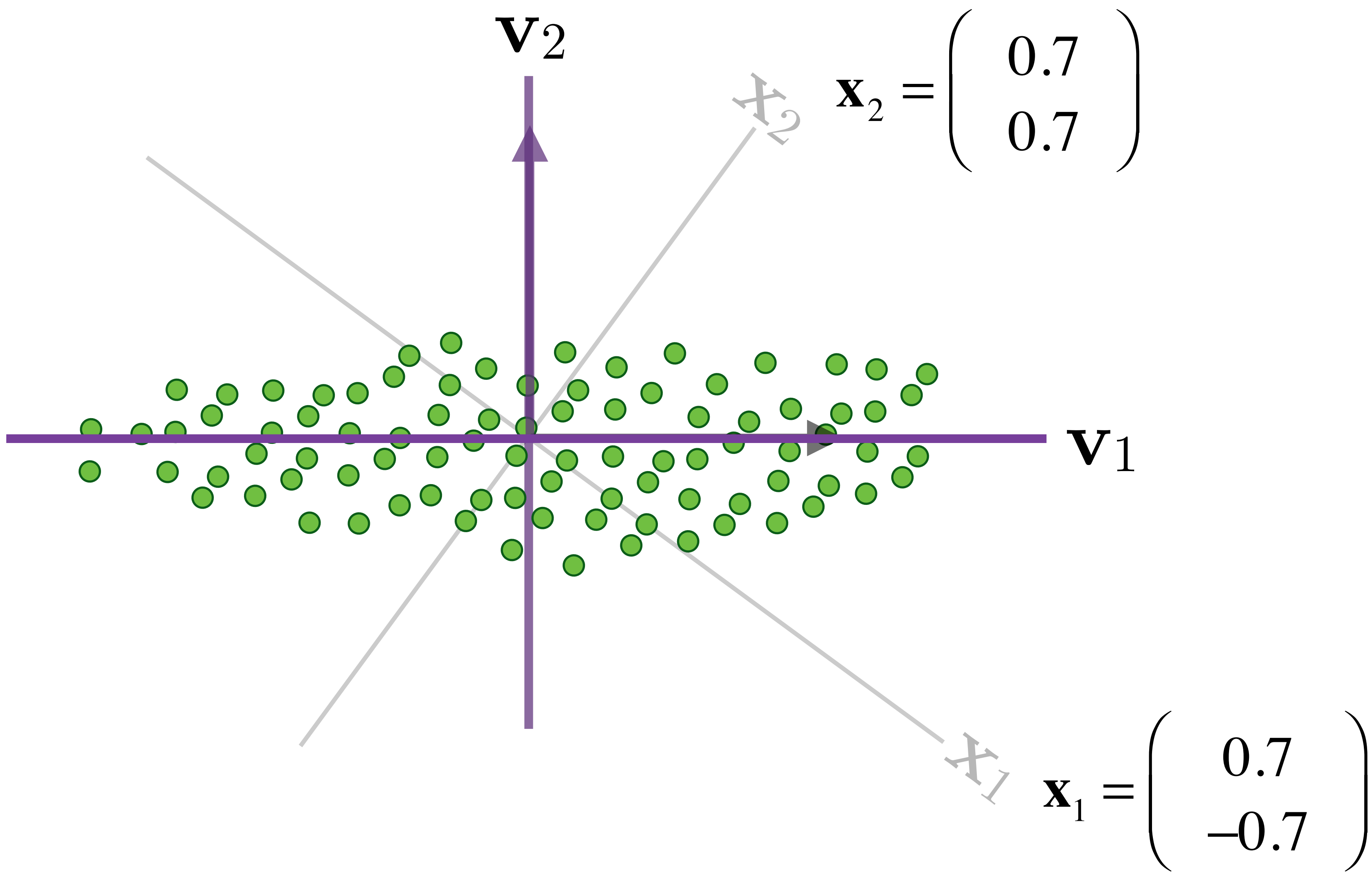
$$\mathbf{v}_1 = \begin{pmatrix} 0.7 \\ 0.7 \end{pmatrix} = 0.7\mathbf{x}_1 + 0.7\mathbf{x}_2$$

$$\mathbf{v}_2 = \begin{pmatrix} -0.7 \\ 0.7 \end{pmatrix} = -0.7\mathbf{x}_1 + 0.7\mathbf{x}_2$$

- These coefficients are called **loadings**

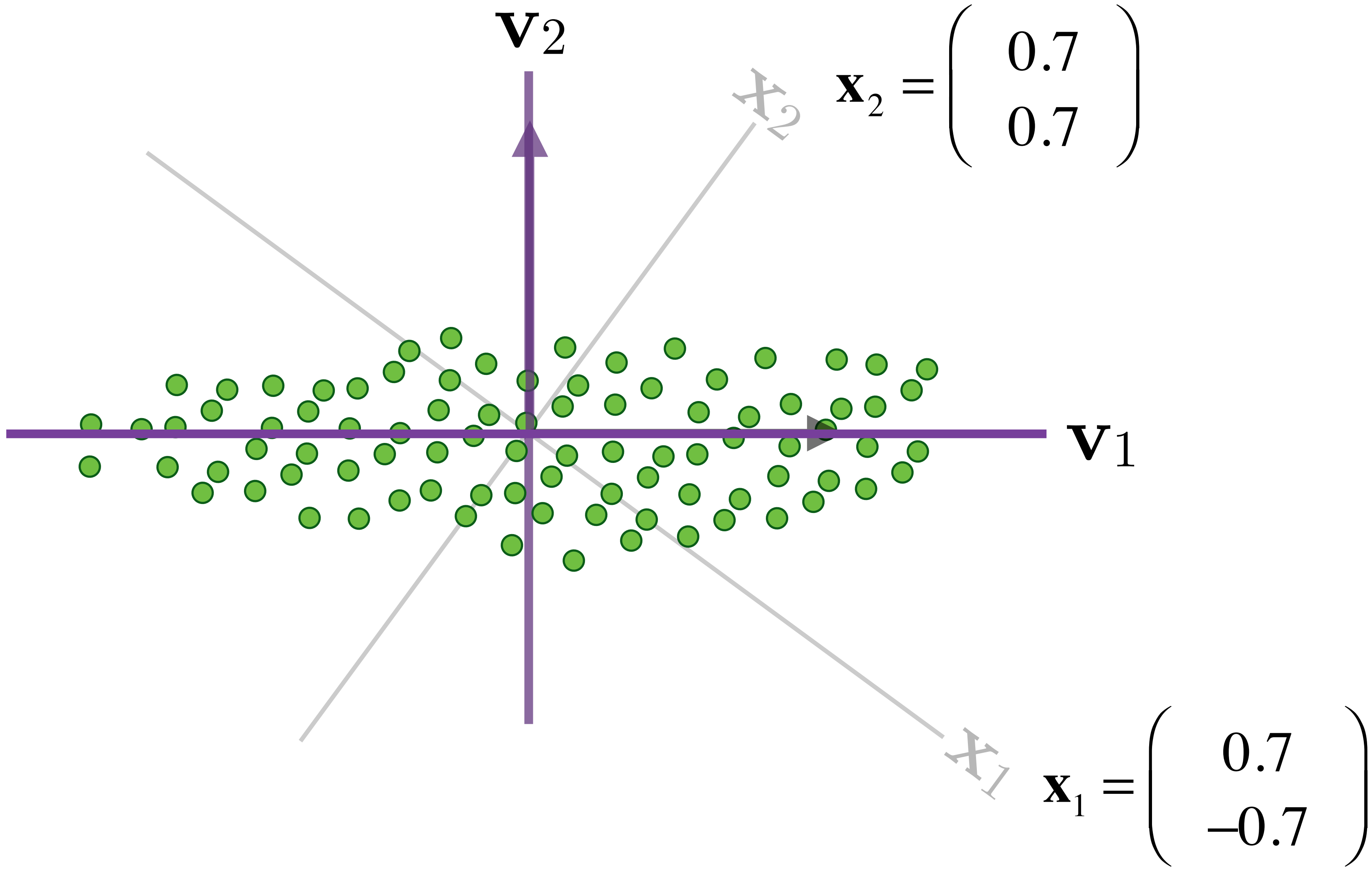
BiPlot

Likewise, we can think of our original basis as linear combinations of the principal components, having coordinates in the new basis!



BiPlot

- Uncorrelated data and variable vectors plotted on same new axes!
- Points in top right have largest x_2 values
- Points in top left have smallest x_1 values



Scores/Coordinates

- The variable loadings give us a formula to compute the coordinates of our data in the new basis.

$$\mathbf{v}_1 = \begin{pmatrix} 0.7 \\ 0.7 \end{pmatrix} = 0.7\mathbf{x}_1 + 0.7\mathbf{x}_2$$

$$\mathbf{v}_2 = \begin{pmatrix} -0.7 \\ 0.7 \end{pmatrix} = -0.7\mathbf{x}_1 + 0.7\mathbf{x}_2$$

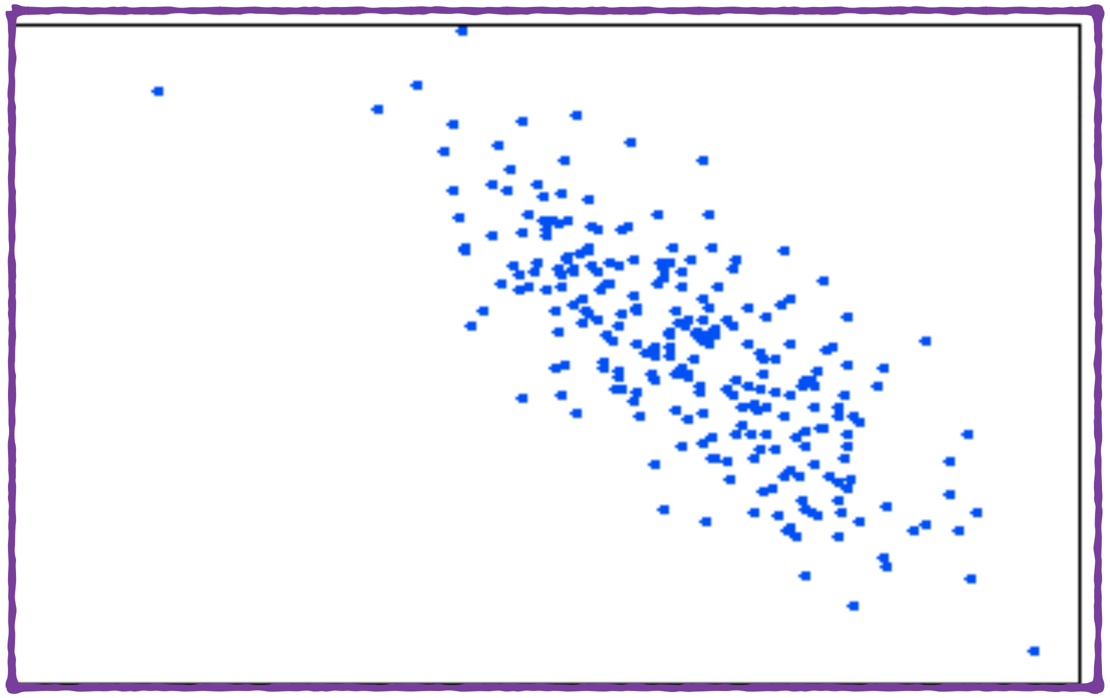
Computing these for each observation gives the new coordinates along the axes $\mathbf{v}_1$ and $\mathbf{v}_2$

- Note that we have to use either the centered data (covariance PCA) or the standardized data (correlation PCA) when using these formulas.

Practice

1

For the following data plot, take your best guess and draw the direction vector for the first and second principal components.



2

Is there more than one correct answer to this question?

Practice

3

Suppose your data contained the 3 variables *VO2.max*, *mile pace*, and *weight* in that order. The first principal component for this data is the eigenvector of the covariance matrix:

$$\begin{pmatrix} 0.69 \\ 0.61 \\ -0.38 \end{pmatrix}$$

What would be the sign of the PC_1 coordinate of an individual with below average *VO2.max*, below average *mile pace*, and above average *weight*?

Major Ideas from Section

- Eigenvector
- Eigenvalue
- Eigenvalue Ordering
- Principal Components
- Directional Variance
- Biplot
- Zero Eigenvalues
- Eigenvectors of symmetric matrices