Principle Components Analysis

Inner Products

Row vector $\mathbf{a} = (a_1, a_2 \ldots, a_n)$

Column vector $\mathbf{b}^T = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{pmatrix}$

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i = \mathbf{a}\mathbf{b}^T = (a_1, a_2 \ldots, a_n) \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{pmatrix}$$

Orthogonality and norm

$\mathbf{a} \cdot \mathbf{b} = 0$ a and b are orthogonal vectors

Norm $\|\mathbf{a}\|_2^2 \doteq \mathbf{a} \cdot \mathbf{a} = \sum_{i=1} a_i^2$

Unit vector $\|\mathbf{a}\|_2 = 1$

matrix-vector product

$$A = \begin{pmatrix} a_{1,1}, a_{1,2}, \dots, a_{1,m} \\ a_{2,1}, a_{2,2}, \dots, a_{2,m} \\ \dots \\ \dots \\ a_{n,1}, a_{n,2}, \dots, a_{n,m} \end{pmatrix} = \begin{pmatrix} \mathbf{a_1} \\ \mathbf{a_2} \\ . \\ . \\ . \\ \mathbf{a_n} \end{pmatrix}$$

$$A\mathbf{b}^T = \begin{pmatrix} \mathbf{a_1}\mathbf{b}^T \\ \mathbf{a_2}\mathbf{b}^T \\ \dots \\ \dots \\ \mathbf{a_n}\mathbf{b}^T \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m a_{1,i}b_i \\ \sum_{i=1}^m a_{2,i}b_i \\ \dots \\ \dots \\ \sum_{i=1}^m a_{n,i}b_i \end{pmatrix}$$

matrix-matrix product

$$A = \begin{pmatrix} a_{1,1}, a_{1,2}, \dots, a_{1,m} \\ a_{2,1}, a_{2,2}, \dots, a_{2,m} \\ \dots \\ \dots \\ a_{n,1}, a_{n,2}, \dots, a_{n,m} \end{pmatrix} = \begin{pmatrix} \mathbf{a_1} \\ \mathbf{a_2} \\ . \\ . \\ . \\ \mathbf{a_n} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{1,1}, b_{1,2}, \dots, b_{1,l} \\ b_{2,1}, b_{2,2}, \dots, b_{2,l} \\ \dots \\ \dots \\ b_{m,1}, b_{m,2}, \dots, b_{m,l} \end{pmatrix} = \left(\mathbf{b_1}^T, \mathbf{b_2}^T, \dots, \mathbf{b_l}^T\right)$$

$$AB = \begin{pmatrix} (\mathbf{a_1} \cdot \mathbf{b_1}), (\mathbf{a_1} \cdot \mathbf{b_2}), \dots, (\mathbf{a_1} \cdot \mathbf{b_l}) \\ (\mathbf{a_2} \cdot \mathbf{b_1}), (\mathbf{a_2} \cdot \mathbf{b_2}), \dots, (\mathbf{a_2} \cdot \mathbf{b_l}) \\ \dots \\ \dots \\ (\mathbf{a_n} \cdot \mathbf{b_1}), (\mathbf{a_n} \cdot \mathbf{b_2}), \dots, (\mathbf{a_n} \cdot \mathbf{b_l}) \end{pmatrix}$$

diagonal matrices

$$D\mathbf{b}^T = \begin{pmatrix} \lambda_1, 0, 0, \ldots, 0 \\ 0, \lambda_2, 0, \ldots, 0 \\ 0, 0, \lambda_3, \ldots, 0 \\ \ldots \\ \ldots \\ 0, 0, 0, \ldots, \lambda_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{pmatrix} = \begin{pmatrix} \lambda_1 b_1 \\ \lambda_2 b_2 \\ . \\ . \\ . \\ \lambda_n b_n \end{pmatrix}$$

$$I\mathbf{b}^T = \begin{pmatrix} 1, 0, 0, \ldots, 0 \\ 0, 1, 0, \ldots, 0 \\ 0, 0, 1, \ldots, 0 \\ \ldots \\ \ldots \\ 0, 0, 0, \ldots, 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_n \end{pmatrix}$$

Orthonormal Matrices

A is a square matrix (nxn)

$$A = \begin{pmatrix} \mathbf{a_1} \\ \mathbf{a_2} \\ . \\ . \\ . \\ \mathbf{a_n} \end{pmatrix} \qquad A^T = (\mathbf{a_1}^T, \mathbf{a_2}^T, \ldots, \mathbf{a_n}^T)$$

$$AA^T = I$$

The rows of ***A*** define an orthonormal basis

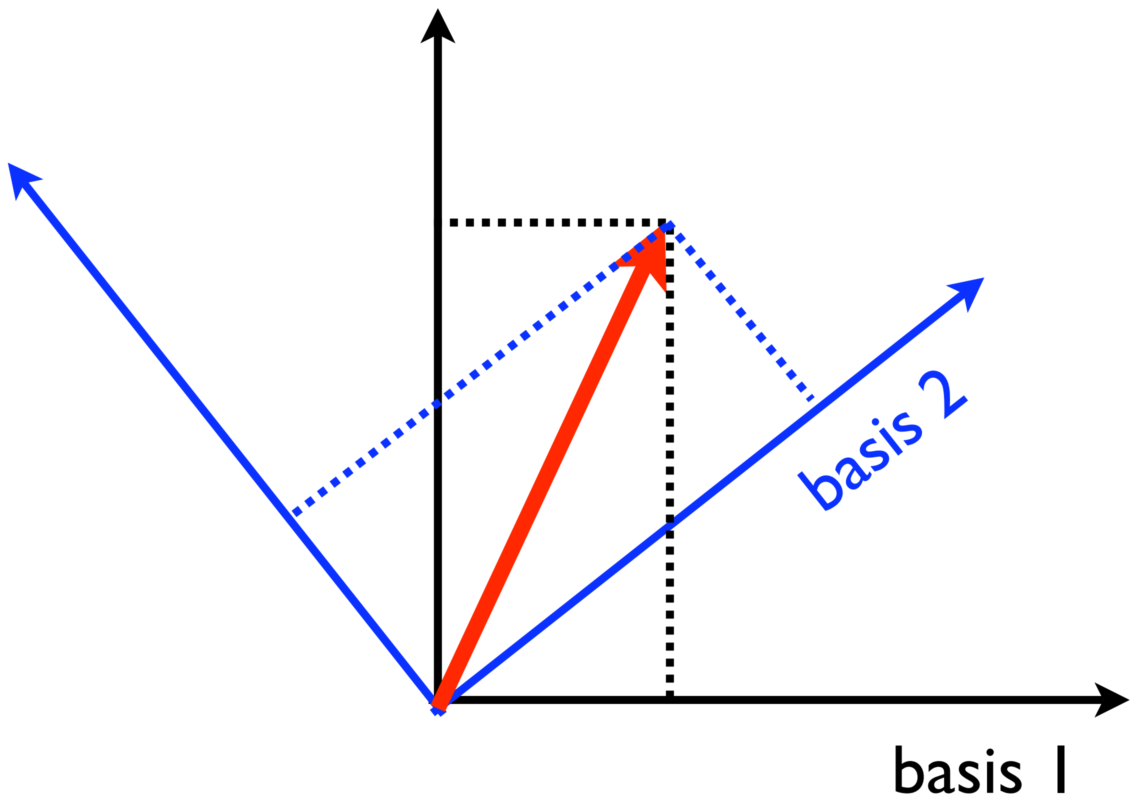
$$\forall i \neq j,\ \mathbf{a_i} \cdot \mathbf{a_i} = \|\mathbf{a}\|_2^2 = 1,\ \mathbf{a_i} \cdot \mathbf{a_j} = 0$$

Orthonormal matrices and vector bases

- A set of n orthogonal unit vectors in R^n defines an orthonormal basis.
- Mutiplying a vector by an orthonormal matrix corresponds to expressing it in terms of the orthonormal basis.

Changing basis in R^2

basis = coordinate system



Eigenvectors and Eigenvalues

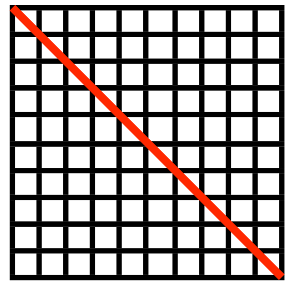
the vector **a** is an eigenvector of the matrix **M** with eigenvalue λ if

$$\mathbf{M}\mathbf{a} = \lambda\mathbf{a}$$

In words: the application of **M** to **a** amounts to changing the length of **a** by a factor of λ without changing **a**'s direction

Decomposing Symmetric Matrices

A symmetic matrix **M**=



M can be written in the form

$$M = A^T \begin{pmatrix} \lambda_1, 0, 0, \ldots, 0 \\ 0, \lambda_2, 0, \ldots, 0 \\ 0, 0, \lambda_3, \ldots, 0 \\ \cdots \\ \cdots \\ 0, 0, 0, \ldots, \lambda_n \end{pmatrix} A$$

A is an orthonormal matrix consisting of the eigenvectors of **M**

Interpretation of the decomposition

- the operation $A^T D A \mathbf{a}$ equals:

  1. transform $\mathbf{a}$ into basis defined by A
  2. multiply each coordinate i by λ_i
  3. transform back to original basis

Principle components of covariance matrices

$$\mathrm{cov}(\mathbf{x}) = \begin{pmatrix} \mathrm{var}(x_1), & \mathrm{cov}(x_1, x_2), \ldots & \mathrm{cov}(x_1, x_n) \\ \mathrm{cov}(x_2, x_1), & \mathrm{var}(x_2), \ldots & \mathrm{cov}(x_2, x_n) \\ & \ldots & \\ & \ldots & \\ \mathrm{cov}(x_n, x_1), & \mathrm{cov}(x_n, x_2), \ldots & \mathrm{var}(x_n) \end{pmatrix}$$

Symmetric matrix because $\mathrm{cov}(x_i, x_j) = \mathrm{cov}(x_j, x_i)$

Can be expressed as cov(x) = $A^T DA$

Interpretation of principle components

- distribution of random vector $\mathbf{x}$ defines covariance matrix $cov(\mathbf{x})$
- decomposition $cov(\mathbf{x}) = A^T D A$ changes the coordinate system of $\mathbf{x}$
- defines a new random vector $\mathbf{y}=A\mathbf{x}$
- $cov(\mathbf{y}) = D$
- $var(y_i)= \lambda_i$, $\forall i \neq j$: $cov(y_i,y_j)=0$