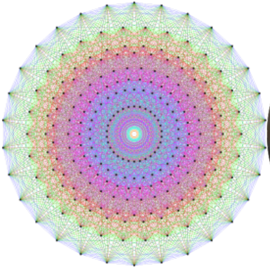
Orthogonal sets and orthogonal matrices

4433LALG3: Linear Algebra

Week 2, Lecture 8, Valente Ramírez

Mathematical & Statistical Methods group — Biometris, Wageningen University & Research



Overview

- Orthogonal and orthonormal sets
- Orthogonal matrices
- Supplementary material

References:

- Nicholson §5.3 and §8.1 (only the subsection *Projections*)

Section 1

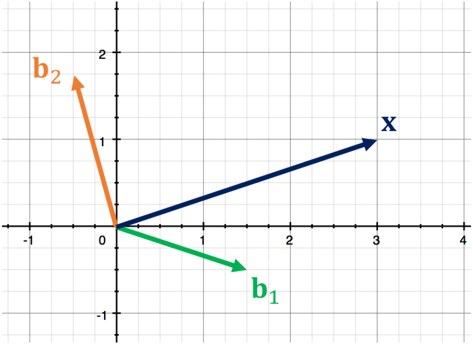
Orthogonal and orthonormal sets

Introduction

Suppose we are in $\mathbb{R}^2$. We have seen that if $\{\mathbf{b}_1, \mathbf{b}_2\}$ is linearly independent, then it constitutes a basis for $\mathbb{R}^2$:

Any vector $\mathbf{x}$ *can be uniquely be expressed as:* $\mathbf{x} = \alpha\mathbf{b}_1 + \beta\mathbf{b}_2$.

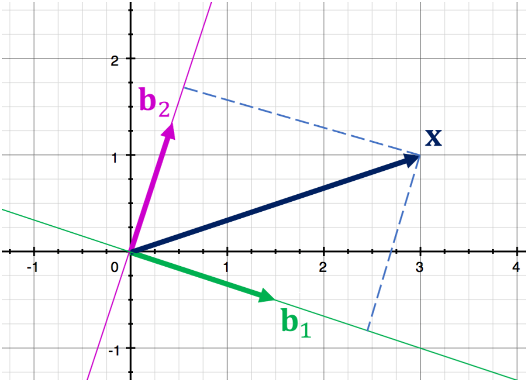
However, decomposing $\mathbf{x}$ in terms of $\mathbf{b}_1, \mathbf{b}_2$ is not always intuitive.



Notice: $\mathbf{x} \neq \mathrm{proj}_{\mathbf{b}_1}\mathbf{x} + \mathrm{proj}_{\mathbf{b}_2}\mathbf{x}$.

Introduction

A very special case is when the vectors $\mathbf{b}_1, \mathbf{b}_2$ are **orthogonal**.



Notice: $\mathbf{x} = \operatorname{proj}_{\mathbf{b}_1} \mathbf{x} + \operatorname{proj}_{\mathbf{b}_2} \mathbf{x}$.

Orthogonal and orthonormal sets

Definition

A set of vectors $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ in $\mathbb{R}^n$ is called an **orthogonal set** if:

- $\mathbf{f}_i \neq \mathbf{0}$ for all $i = 1, \ldots, k$,
- $\mathbf{f}_i^\top \mathbf{f}_j = 0$ whenever $i \neq j$.

Moreover, an orthogonal set is called **orthonormal** if it also satisfies:

- $\|\mathbf{f}_i\| = 1$ for all $i = 1, \ldots, k$.

A common way of expressing that a set is orthonormal is by writing:

$$\mathbf{f}_i^\top \mathbf{f}_j = \delta_{ij},$$

where δ_{ij} is the *Kronecker delta* function.

Some important facts

Definition

We say that a set $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ is an **orthogonal basis** for a subspace U of $\mathbb{R}^n$ whenever the set is an orthogonal set and it is also a basis for U.

Theorem

1. *Any orthogonal set* $\{\mathbf{f}_1, \ldots, \mathbf{f}_k\}$ *in* $\mathbb{R}^n$ *is linearly independent.*
2. *Any subspace* U *of* $\mathbb{R}^n$ *has an orthogonal basis.*
3. *Any orthogonal set in* $\mathbb{R}^n$ *of size* n *is an orthogonal basis for* $\mathbb{R}^n$*.*

In the above definition and theorem, all instances of *'orthogonal'* could be replaced by *'orthonormal'* and the results remain valid.

Expansion theorem

Here we formalize our claim on slide 4.

Theorem

Let $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\}$ *be an orthogonal basis of a subspace* U *of* $\mathbb{R}^n$*.*

Let $\mathbf{y}$ *be an arbitrary vector in* $\mathbb{R}^n$*. Then:*

$$\operatorname{proj}_U \mathbf{y} = \operatorname{proj}_{\mathbf{f}_1} \mathbf{y} + \operatorname{proj}_{\mathbf{f}_2} \mathbf{y} + \ldots + \operatorname{proj}_{\mathbf{f}_k} \mathbf{y}.$$

In particular,[1] *if* $\mathbf{x}$ *is a vector in* U*, then:*

$$\mathbf{x} = \operatorname{proj}_{\mathbf{f}_1} \mathbf{x} + \operatorname{proj}_{\mathbf{f}_2} \mathbf{x} + \ldots + \operatorname{proj}_{\mathbf{f}_k} \mathbf{x}.$$

[1] This part is the *expansion theorem* in Nicholson (Theorem 5.3.6). You can recover his formula by using the following fact from last lecture: $\operatorname{proj}_{\mathbf{f}} \mathbf{x} = \frac{\mathbf{f}^\top \mathbf{x}}{\|\mathbf{f}\|^2} \mathbf{f}$.

The orthonormal case

We will focus on the case of an orthonormal set $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$.

From last lecture, we know that if $\mathbf{u}$ is of unit length, then:

$$\mathrm{proj}_{\mathbf{u}}\,\mathbf{x} = (\mathbf{u}^\top \mathbf{x})\,\mathbf{u}.$$

We conclude that:

$$\mathrm{proj}_U\,\mathbf{y} = (\mathbf{u}_1^\top \mathbf{y})\,\mathbf{u}_1 + \ldots + (\mathbf{u}_k^\top \mathbf{y})\,\mathbf{u}_k.$$

The orthonormal case

We now re-write the last formula using matrix notation.

So far, we have been treating $\mathbf{u}_i^\top \mathbf{y}$ as a scalar (because it is). But we can also think of it as a 1×1 matrix.

If we want to multiply it against the $n \times 1$ vector $\mathbf{u}_i$, we need to do the multiplication from the right:

$$\begin{aligned} \operatorname{proj}_U \mathbf{y} &= (\mathbf{u}_1^\top \mathbf{y})\,\mathbf{u}_1 + \ldots + (\mathbf{u}_k^\top \mathbf{y})\,\mathbf{u}_k \\ &= \mathbf{u}_1(\mathbf{u}_1^\top \mathbf{y}) + \ldots + \mathbf{u}_k(\mathbf{u}_k^\top \mathbf{y}) \\ &= (\mathbf{u}_1\mathbf{u}_1^\top)\mathbf{y} + \ldots + (\mathbf{u}_k\mathbf{u}_k^\top)\mathbf{y} \end{aligned}$$

We conclude the following:

Useful facts

Let P be the matrix associated to the orthogonal projection onto U, and for $i = 1, \ldots, k$, let P_i be the matrix associated to projection onto $\mathbf{u}_i$.

- $P = P_1 + \ldots + P_k$,
- $P_i = \mathbf{u}_i\mathbf{u}_i^\top$.

Section 2

Orthogonal matrices

Orthogonal matrices

An important situation is when we have an orthonormal basis for $\mathbb{R}^n$. In this case, it is usual to arrange the vectors into an $n \times n$ matrix.

Definition

An $n \times n$ matrix P is called an **orthogonal matrix** whenever the columns of P form an orthonormal set.

Warning

The columns of an orthogonal matrix are required to be orthonormal, not just orthogonal!

Orthogonal matrices

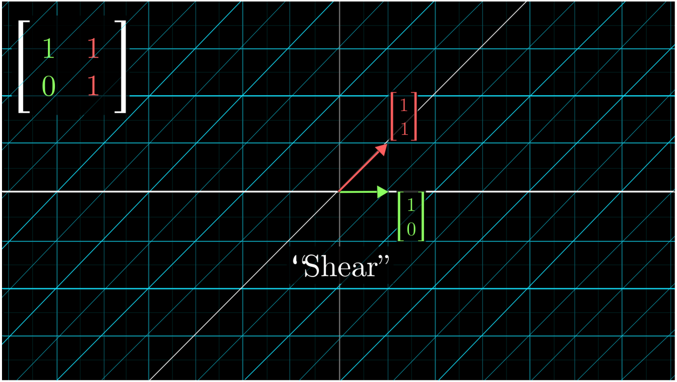
Properties

Suppose P is an $n \times n$ orthogonal matrix. Then the following hold:

- $P^\top$ is also orthogonal
  - In particular, the **rows** of P also from an orthonormal set
- $\det P = \pm 1$
  - In particular, P is invertible
- $P^{-1} = P^\top$

Distortion

Usually, linear transformations distort both length and angles.



No distortion

Suppose T is a transformation that does **not** distort lengths nor angles.

What can we say about the columns of the associated matrix A?

- Because $\mathbf{a}_i = T(\mathbf{e}_i)$, we have $\|\mathbf{a}_i\| = \|\mathbf{e}_i\| = 1$.
- Because $\mathbf{e}_i, \mathbf{e}_j$ are orthogonal for $i \neq j$, so are $\mathbf{a}_i, \mathbf{a}_j$.

Conclusion:

The matrix A is orthogonal.

Orthogonal transformations

A linear transformation is called **orthogonal** (or sometimes *an isometry*) if it is represented by an orthogonal matrix.

Orthogonal transformations are precisely those linear transformations that **do not distort** angles and distances.

Indeed, computing the dot product before or after applying the transformation gives the same result.

If P is an orthogonal matrix:

$$\begin{aligned}(P\mathbf{v})^\top(P\mathbf{w}) &= (\mathbf{v}^\top P^\top)(P\mathbf{w})\\ &= \mathbf{v}^\top(P^\top P)\mathbf{w}\\ &= \mathbf{v}^\top I_n\mathbf{w}\\ &= \mathbf{v}^\top\mathbf{w}\end{aligned}$$

Section 3

Supplementary material

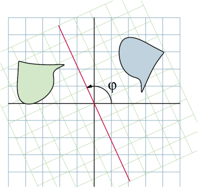
Rotations in $\mathbb{R}^2$

Example

A (counter-clockwise) rotation about the origin with angle φ preserves the length of any vector, and the angle between any two vectors.
Therefore, the associated matrix P must be orthogonal.

With some trigonometry, it is easy to derive that:

$$P = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}.$$



Exercise:

Verify that the columns of P form an orthonormal set, and compute the determinant.

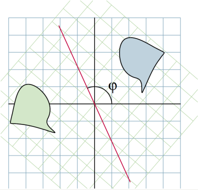
Reflections in $\mathbb{R}^2$

Example

Suppose ℓ is the line through the origin forming an angle φ with the x–axis. The reflection in the line ℓ defines a linear transformation which preserves the length of any vector, and the angle between any two vectors (disregarding orientation). Therefore, the associated matrix Q must be orthogonal.

With some trigonometry, it is easy to derive that:

$$Q = \begin{bmatrix} \cos(2\varphi) & \sin(2\varphi) \\ \sin(2\varphi) & -\cos(2\varphi) \end{bmatrix}.$$



Exercise:

Verify that the columns of Q form an orthonormal set, and compute the determinant.

Orthogonal matrices

The product of two orthogonal matrices is also orthogonal.

The simplest examples of orthogonal matrices are *rotations* and *reflections*. Therefore, also the composition of such transformations results in an orthogonal transformation.

Exercise:

Consider $P = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, which represents a 90° rotation clockwise, and $Q = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, which represents reflection in the y–axis.

Compute the matrix associated to the transformation "90° rotation clockwise followed by reflection in the y–axis", and verify that this is an orthogonal matrix.

Is this a rotation or a reflection?