Math 415 - Lecture 25

Multiple linear regression, Gram Schmidt and Orthogonal matrices

Monday October 26th 2015

Textbook reading: Chapters 3.3,3.4

Textbook reading: Chapters 3.3,3.4

Suggested practice exercises: Chapter 3.3, 3,5,6,13,22,24,25,26
and Chapter 3.4, 10,11,13,14,16,26

Textbook reading: Chapters 3.3,3.4

Suggested practice exercises: Chapter 3.3, 3,5,6,13,22,24,25,26 and Chapter 3.4, 10,11,13,14,16,26

Khan Academy video: Another Least Squares Example, Gram-Schmidt Example

Textbook reading: Chapters 3.3,3.4

Suggested practice exercises: Chapter 3.3, 3,5,6,13,22,24,25,26 and Chapter 3.4, 10,11,13,14,16,26

Khan Academy video: Another Least Squares Example, Gram-Schmidt Example

Strang lecture: Orthogonal Matrices and Gram-Schmidt Process.

Review

$\hat{\mathbf{x}}$ is a least squares solution of the system $A\mathbf{x} = \mathbf{b}$

$\hat{\mathbf{x}}$ is a least squares solution of the system $A\mathbf{x} = \mathbf{b}$

$\iff$ $\hat{\mathbf{x}}$ is such that $A\hat{\mathbf{x}} - \mathbf{b}$ is as small as possible

$\hat{\mathbf{x}}$ is a least squares solution of the system $A\mathbf{x} = \mathbf{b}$

$\iff$ $\hat{\mathbf{x}}$ is such that $A\hat{\mathbf{x}} - \mathbf{b}$ is as small as possible

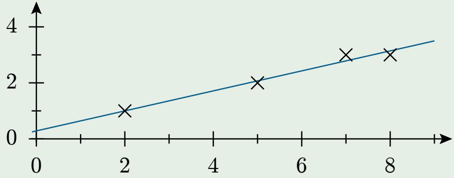
$\overset{FTLA}{\iff}$ $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$ (the **normal equations**)

Application: fitting data

Least square lines

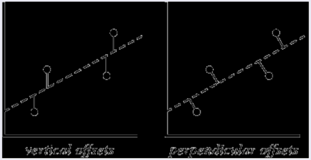
Example

Find β_1, β_2 such that the line $y = \beta_1 + \beta_2 x$ best fits the data points $(2,1), (5,2), (7,3), (8,3)$.



Comment

As usual in practice, we are minimizing the (the sum of the squares of the) vertical offsets.



Solution

Solution

The equations $y = \beta_1 + \beta_2 x$ in matrix form:

Solution

The equations $y = \beta_1 + \beta_2 x$ in matrix form:

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\text{observation vector } \mathbf{y}}$$

The equations $y = \beta_1 + \beta_2 x$ in matrix form:

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\text{observation vector } \mathbf{y}}$$

Here, we need to find a least squares solution to

Solution

The equations $y = \beta_1 + \beta_2 x$ in matrix form:

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\text{observation vector } \mathbf{y}}$$

Here, we need to find a least squares solution to

$$\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}.$$

Solution

$$X^T X =$$

Solution

$$X^TX = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

Solution

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T \mathbf{y} =$$

Solution

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Solution

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Solving $\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \hat{\beta} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$, we find $\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} =$

Solution

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Solving $\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \hat{\beta} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$, we find $\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$.

Solution

$$X^TX = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T\mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Solving $\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \hat{\beta} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$, we find $\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$.

Hence the least squares line is

Solution

$$X^TX = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T\mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Solving $\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \hat{\beta} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$, we find $\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$.

Hence the least squares line is $y = \frac{2}{7} + \frac{5}{14}x$.

Fitting to other curves

What happens if the data just lie close to any line? We can also fit the experimental data using other curves. Try to find $\beta_1, \beta_2, \beta_3$ such that $y = \beta_1 + \beta_2 x + \beta_3 x^2$ fits the data.

What happens if the data just lie close to any line? We can also fit the experimental data using other curves. Try to find $\beta_1, \beta_2, \beta_3$ such that $y = \beta_1 + \beta_2 x + \beta_3 x^2$ fits the data. To fit: $y_i \approx \beta_1 + \beta_2 x_i + \beta_3 x_i^2$ with parameters $\beta_1, \beta_2, \beta_3$.

What happens if the data just lie close to any line? We can also fit the experimental data using other curves. Try to find $\beta_1, \beta_2, \beta_3$ such that $y = \beta_1 + \beta_2 x + \beta_3 x^2$ fits the data. To fit:
$y_i \approx \beta_1 + \beta_2 x_i + \beta_3 x_i^2$ with parameters $\beta_1, \beta_2, \beta_3$.
The equations $y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2$ in matrix form:

What happens if the data just lie close to any line? We can also fit the experimental data using other curves. Try to find $\beta_1, \beta_2, \beta_3$ such that $y = \beta_1 + \beta_2 x + \beta_3 x^2$ fits the data. To fit: $y_i \approx \beta_1 + \beta_2 x_i + \beta_3 x_i^2$ with parameters $\beta_1, \beta_2, \beta_3$.
The equations $y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2$ in matrix form:

$$\underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ \vdots & \vdots & \vdots \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \end{bmatrix}}_{\text{observation vector } \mathbf{y}}$$

What happens if the data just lie close to any line? We can also fit the experimental data using other curves. Try to find $\beta_1, \beta_2, \beta_3$ such that $y = \beta_1 + \beta_2 x + \beta_3 x^2$ fits the data. To fit: $y_i \approx \beta_1 + \beta_2 x_i + \beta_3 x_i^2$ with parameters $\beta_1, \beta_2, \beta_3$.
The equations $y_i = \beta_1 + \beta_2 x_i + \beta_3 x_i^2$ in matrix form:

$$\underbrace{\begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ \vdots & \vdots & \vdots \end{bmatrix}}_{\text{design matrix } X} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \end{bmatrix}}_{\text{observation vector } \mathbf{y}}$$

Given data (x_i, y_i), we then find the least squares solution to $X\beta = \mathbf{y}$.

Multiple linear regression

Of course, sometimes the variable y might not just depend on a single variable x, but on two variables, say u and v. So, here you have find the least-squares solution of

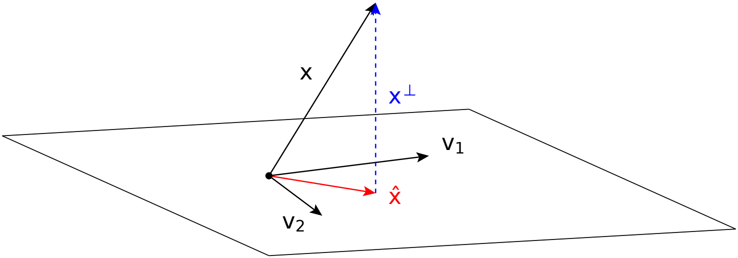
$$\underbrace{\begin{bmatrix} 1 & v_1 & w_1 \\ 1 & v_2 & w_2 \\ 1 & v_3 & w_3 \\ \vdots & \vdots & \vdots \end{bmatrix}}_{\text{design matrix}} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \end{bmatrix}}_{\text{observation vector}}$$

Of course, sometimes the variable y might not just depend on a single variable x, but on two variables, say u and v. So, here you have find the least-squares solution of

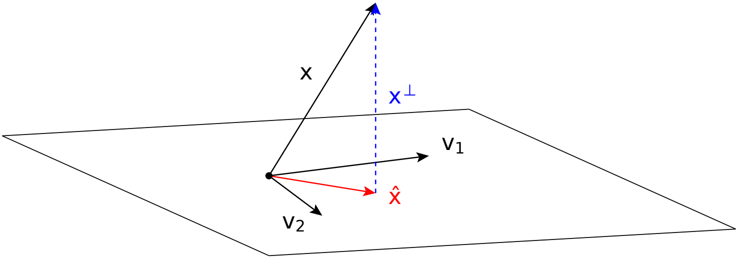
$$\underbrace{\begin{bmatrix} 1 & v_1 & w_1 \\ 1 & v_2 & w_2 \\ 1 & v_3 & w_3 \\ \vdots & \vdots & \vdots \end{bmatrix}}_{\text{design matrix}} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \end{bmatrix}}_{\text{observation vector}}$$

And we again proceed by finding a least squares solution.

Review

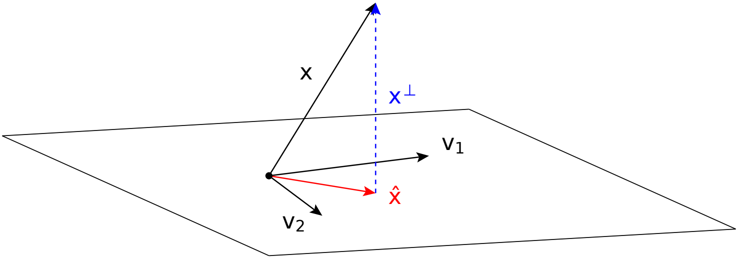


Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is an orthonormal basis of W. The **orthogonal projection** of $\mathbf{x}$ onto W is :



Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is an orthonormal basis of W. The **orthogonal projection** of $\mathbf{x}$ onto W is :

$$\hat{\mathbf{x}} = \underbrace{\langle \mathbf{x}, \mathbf{v}_1 \rangle \mathbf{v}_1}_{\text{proj. of } \mathbf{x} \text{ onto } \mathbf{v}_1} + \ldots + \underbrace{\langle \mathbf{x}, \mathbf{v}_m \rangle \mathbf{v}_m}_{\text{proj. of } \mathbf{x} \text{ onto } \mathbf{v}_m}$$



Suppose $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is an orthonormal basis of W. The **orthogonal projection** of $\mathbf{x}$ onto W is :

$$\hat{\mathbf{x}} = \underbrace{\langle \mathbf{x}, \mathbf{v}_1 \rangle \mathbf{v}_1}_{\text{proj. of } \mathbf{x} \text{ onto } \mathbf{v}_1} + \ldots + \underbrace{\langle \mathbf{x}, \mathbf{v}_m \rangle \mathbf{v}_m}_{\text{proj. of } \mathbf{x} \text{ onto } \mathbf{v}_m}$$

(To stay agile, we are writing $\langle \mathbf{x}, \mathbf{v}_1 \rangle = \mathbf{x} \cdot \mathbf{v}_1$ for the inner product.)

Gram-Schmidt

Our goal

* In calculating projections we used an *orthogonal basis* and the easy formula for the coefficients.

* In calculating projections we used an *orthogonal basis* and the easy formula for the coefficients.
* What if we are given an arbitrary basis, not orthogonal?

* In calculating projections we used an *orthogonal basis* and the easy formula for the coefficients.
* What if we are given an arbitrary basis, not orthogonal?
* Turn the starting basis into an orthogonal (or orthonormal) basis.

Our goal

* In calculating projections we used an *orthogonal basis* and the easy formula for the coefficients.
* What if we are given an arbitrary basis, not orthogonal?
* Turn the starting basis into an orthogonal (or orthonormal) basis.
* **Gram-Schmidt Process.**

Our goal

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

$$\mathbf{b}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}$$

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

$$\mathbf{b}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}$$

$$\mathbf{b}_2 = \mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{q}_1 \rangle \mathbf{q}_1, \qquad \mathbf{q}_2 = \frac{\mathbf{b}_2}{\|\mathbf{b}_2\|}$$

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \dots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \dots, \mathbf{q}_n$.

$$\begin{aligned}
\mathbf{b}_1 &= \mathbf{a}_1, & \mathbf{q}_1 &= \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|} \\
\mathbf{b}_2 &= \mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{q}_1 \rangle \mathbf{q}_1, & \mathbf{q}_2 &= \frac{\mathbf{b}_2}{\|\mathbf{b}_2\|} \\
\mathbf{b}_3 &= \mathbf{a}_3 - \langle \mathbf{a}_3, \mathbf{q}_1 \rangle \mathbf{q}_1 - \langle \mathbf{a}_3, \mathbf{q}_2 \rangle \mathbf{q}_2, & \mathbf{q}_3 &= \frac{\mathbf{b}_3}{\|\mathbf{b}_3\|} \\
&\dots & &\dots
\end{aligned}$$

Example

Find an orthonormal basis for $V = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Solution

$$\mathbf{b}_1 =$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{b}_2 =$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \left\langle \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_1 \right\rangle \mathbf{q}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

$$\mathbf{b}_2 = \begin{bmatrix}2\\1\\0\\0\end{bmatrix} - \langle \begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \mathbf{q}_1 \rangle \mathbf{q}_1 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \qquad \mathbf{q}_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \langle \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{q}_1 \rangle \mathbf{q}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{q}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{b}_3 =$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

$$\mathbf{b}_2 = \begin{bmatrix}2\\1\\0\\0\end{bmatrix} - \left\langle \begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \mathbf{q}_1 \right\rangle \mathbf{q}_1 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \qquad \mathbf{q}_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

$$\mathbf{b}_3 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} -$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

$$\mathbf{b}_2 = \begin{bmatrix}2\\1\\0\\0\end{bmatrix} - \left\langle \begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \mathbf{q}_1 \right\rangle \mathbf{q}_1 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \qquad \mathbf{q}_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

$$\mathbf{b}_3 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \left\langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{q}_1 \right\rangle \mathbf{q}_1 -$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

$$\mathbf{b}_2 = \begin{bmatrix}2\\1\\0\\0\end{bmatrix} - \langle \begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \mathbf{q}_1\rangle \mathbf{q}_1 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \qquad \mathbf{q}_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

$$\mathbf{b}_3 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{q}_1\rangle \mathbf{q}_1 - \langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{q}_2\rangle \mathbf{q}_2 =$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

$$\mathbf{b}_2 = \begin{bmatrix}2\\1\\0\\0\end{bmatrix} - \left\langle \begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \mathbf{q}_1\right\rangle \mathbf{q}_1 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \qquad \mathbf{q}_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

$$\mathbf{b}_3 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \left\langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{q}_1\right\rangle \mathbf{q}_1 - \left\langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{q}_2\right\rangle \mathbf{q}_2 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix},$$

Solution

$$\mathbf{b}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \qquad \mathbf{q}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

$$\mathbf{b}_2 = \begin{bmatrix}2\\1\\0\\0\end{bmatrix} - \langle \begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \mathbf{q}_1\rangle \mathbf{q}_1 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \qquad \mathbf{q}_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

$$\mathbf{b}_3 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{q}_1\rangle \mathbf{q}_1 - \langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{q}_2\rangle \mathbf{q}_2 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix}, \quad \mathbf{q}_3 = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\0\\1\\1\end{bmatrix}$$

We have obtained an orthonormal basis for V: $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$.

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Why does Gram-Schmidt work?

Why does Gram-Schmidt work? Recall, if W is a subspace, $\mathbf{b}$ any vector, then

Why does Gram-Schmidt work? Recall, if W is a subspace, $\mathbf{b}$ any vector, then

$$\hat{\mathbf{b}} \rightsquigarrow \text{ projection to } W, \ \mathbf{b}^{\perp} = \mathbf{b} - \hat{\mathbf{b}} \rightsquigarrow \text{ orth. to } W$$

Why does Gram-Schmidt work? Recall, if W is a subspace, $\mathbf{b}$ any vector, then

$$\hat{\mathbf{b}} \rightsquigarrow \text{ projection to } W, \ \mathbf{b}^{\perp} = \mathbf{b} - \hat{\mathbf{b}} \rightsquigarrow \text{ orth. to } W$$

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

Why does Gram-Schmidt work? Recall, if W is a subspace, $\mathbf{b}$ any vector, then

$$\hat{\mathbf{b}} \rightsquigarrow \text{ projection to } W, \ \mathbf{b}^{\perp} = \mathbf{b} - \hat{\mathbf{b}} \rightsquigarrow \text{ orth. to } W$$

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

Why does Gram-Schmidt work? Recall, if W is a subspace, $\mathbf{b}$ any vector, then

$$\hat{\mathbf{b}} \rightsquigarrow \text{ projection to } W, \; \mathbf{b}^{\perp} = \mathbf{b} - \hat{\mathbf{b}} \rightsquigarrow \text{ orth. to } W$$

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

$$\mathbf{b}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \underbrace{\frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}}_{\textit{normalize}}$$

Why does Gram-Schmidt work? Recall, if W is a subspace, $\mathbf{b}$ any vector, then

$$\hat{\mathbf{b}} \rightsquigarrow \text{ projection to } W, \ \mathbf{b}^{\perp} = \mathbf{b} - \hat{\mathbf{b}} \rightsquigarrow \text{ orth. to } W$$

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

$$\mathbf{b}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \underbrace{\frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}}_{\textit{normalize}}$$

$$\mathbf{b}_2 = \underbrace{\mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{q}_1 \rangle \mathbf{q}_1}_{\mathbf{a}_2^{\perp} \rightsquigarrow \text{orth. to Span}\{\mathbf{q}_1\}},$$

Why does Gram-Schmidt work? Recall, if W is a subspace, $\mathbf{b}$ any vector, then

$$\hat{\mathbf{b}} \rightsquigarrow \text{ projection to } W, \ \mathbf{b}^{\perp} = \mathbf{b} - \hat{\mathbf{b}} \rightsquigarrow \text{ orth. to } W$$

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

$$\mathbf{b}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \underbrace{\frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}}_{normalize}$$

$$\mathbf{b}_2 = \underbrace{\mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{q}_1 \rangle \mathbf{q}_1}_{\mathbf{a}_2^{\perp} \rightsquigarrow \text{orth. to Span}\{\mathbf{q}_1\}}, \qquad \mathbf{q}_2 = \underbrace{\frac{\mathbf{b}_2}{\|\mathbf{b}_2\|}}_{normalize}$$

Why does Gram-Schmidt work? Recall, if W is a subspace, $\mathbf{b}$ any vector, then

$$\hat{\mathbf{b}} \rightsquigarrow \text{ projection to } W, \; \mathbf{b}^{\perp} = \mathbf{b} - \hat{\mathbf{b}} \rightsquigarrow \text{ orth. to } W$$

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \ldots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_n$.

$$\mathbf{b}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \underbrace{\frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}}_{\textit{normalize}}$$

$$\mathbf{b}_2 = \underbrace{\mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{q}_1 \rangle \mathbf{q}_1}_{\mathbf{a}_2^{\perp} \rightsquigarrow \text{orth. to Span}\{\mathbf{q}_1\}}, \qquad \mathbf{q}_2 = \underbrace{\frac{\mathbf{b}_2}{\|\mathbf{b}_2\|}}_{\textit{normalize}}$$

$$\mathbf{b}_3 = \underbrace{\mathbf{a}_3 - (\langle \mathbf{a}_3, \mathbf{q}_1 \rangle \mathbf{q}_1 + \langle \mathbf{a}_3, \mathbf{q}_2 \rangle \mathbf{q}_2)}_{\mathbf{a}_3^{\perp} \text{ orth. to Span}\{\mathbf{q}_1, \mathbf{q}_2\}}$$

Why does Gram-Schmidt work? Recall, if W is a subspace, $\mathbf{b}$ any vector, then

$$\hat{\mathbf{b}} \rightsquigarrow \text{ projection to } W,\ \mathbf{b}^{\perp} = \mathbf{b} - \hat{\mathbf{b}} \rightsquigarrow \text{ orth. to } W$$

Recipe. (Gram-Schmidt orthonormalization)

Given a basis $\mathbf{a}_1, \dots, \mathbf{a}_n$, produce a orthogonal basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ and an orthonormal basis $\mathbf{q}_1, \dots, \mathbf{q}_n$.

$$\mathbf{b}_1 = \mathbf{a}_1, \qquad \mathbf{q}_1 = \underbrace{\frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}}_{\textit{normalize}}$$

$$\mathbf{b}_2 = \underbrace{\mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{q}_1 \rangle \mathbf{q}_1}_{\mathbf{a}_2^{\perp} \rightsquigarrow \text{orth. to Span}\{\mathbf{q}_1\}}, \qquad \mathbf{q}_2 = \underbrace{\frac{\mathbf{b}_2}{\|\mathbf{b}_2\|}}_{\textit{normalize}}$$

$$\mathbf{b}_3 = \underbrace{\mathbf{a}_3 - (\langle \mathbf{a}_3, \mathbf{q}_1 \rangle \mathbf{q}_1 + \langle \mathbf{a}_3, \mathbf{q}_2 \rangle \mathbf{q}_2)}_{\mathbf{a}_3^{\perp} \text{ orth. to Span}\{\mathbf{q}_1, \mathbf{q}_2\}} \qquad \underbrace{\mathbf{q}_3 = \frac{\mathbf{b}_3}{\|\mathbf{b}_3\|}}_{\textit{normalize}}$$

Example

Let $V = \text{Span}\{\begin{bmatrix}2\\1\\2\end{bmatrix}, \begin{bmatrix}0\\0\\3\end{bmatrix}\}$. Find an orthonormal basis for V.

Check that your basis is actually orthonormal.

Orthogonal matrices

Theorem

Let $A = \begin{bmatrix}\mathbf{a_1}, \ldots .\mathbf{a_n}\end{bmatrix}$ be ay matrix. Then $A^T A$ is the matrix of dot products of the columns of A:

Theorem

Let $A = [\mathbf{a_1}, \ldots .\mathbf{a_n}]$ be ay matrix. Then $A^T A$ is the matrix of dot products of the columns of A:

Theorem

Let $A = \begin{bmatrix}\mathbf{a_1}, \ldots .\mathbf{a_n}\end{bmatrix}$ be ay matrix. Then A^TA is the matrix of dot products of the columns of A:

$$A^TA = \begin{bmatrix} \mathbf{a_1}\cdot\mathbf{a_1} & \mathbf{a_1}\cdot\mathbf{a_2} & \mathbf{a_1}\cdot\mathbf{a_3} & \ldots \\ \mathbf{a_2}\cdot\mathbf{a_1} & \mathbf{a_2}\cdot\mathbf{a_2} & \mathbf{a_2}\cdot\mathbf{a_3} & \ldots \\ \mathbf{a_3}\cdot\mathbf{a_1} & \mathbf{a_3}\cdot\mathbf{a_2} & \mathbf{a_3}\cdot\mathbf{a_3} & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Theorem

Let $A = \begin{bmatrix} \mathbf{a_1}, \ldots . \mathbf{a_n} \end{bmatrix}$ *be ay matrix. Then* $A^T A$ *is the matrix of dot products of the columns of* A*:*

$$A^T A = \begin{bmatrix} \mathbf{a_1} \cdot \mathbf{a_1} & \mathbf{a_1} \cdot \mathbf{a_2} & \mathbf{a_1} \cdot \mathbf{a_3} & \ldots \\ \mathbf{a_2} \cdot \mathbf{a_1} & \mathbf{a_2} \cdot \mathbf{a_2} & \mathbf{a_2} \cdot \mathbf{a_3} & \ldots \\ \mathbf{a_3} \cdot \mathbf{a_1} & \mathbf{a_3} \cdot \mathbf{a_2} & \mathbf{a_3} \cdot \mathbf{a_3} & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

What happens if the columns of A are orthonormal?

Theorem

The columns of Q are orthonormal $\iff$ $Q^T Q = I$

Theorem

The columns of Q are orthonormal $\iff Q^TQ = I$

Proof.

Let $\mathbf{q}_1, \ldots, \mathbf{q}_n$ be the columns of Q.

They orthonormal if and only if $\mathbf{q}_i^T\mathbf{q}_j = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}$

All these products are packaged in $Q^TQ = I$:

$$\begin{bmatrix} -- & \mathbf{q}_1^T & -- \\ -- & \mathbf{q}_2^T & -- \\ & \vdots & \end{bmatrix} \begin{bmatrix} | & | & \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots \\ | & | & \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \ddots \end{bmatrix}$$

$\square$

Definition

An **orthogonal matrix** is a square matrix Q with orthonormal columns.

Definition

An **orthogonal matrix** is a square matrix Q with orthonormal columns.

It is historical convention to restrict to square matrices, and to say orthogonal matrix even though "orthonormal matrix" might be better.

Definition

An **orthogonal matrix** is a square matrix Q with orthonormal columns.

It is historical convention to restrict to square matrices, and to say orthogonal matrix even though "orthonormal matrix" might be better.

An $n \times n$ matrix Q is orthogonal $\iff Q^TQ = I$
In other words, $Q^{-1} = Q^T$.

Definition

An **orthogonal matrix** is a square matrix Q with orthonormal columns.

It is historical convention to restrict to square matrices, and to say orthogonal matrix even though "orthonormal matrix" might be better.

An $n \times n$ matrix Q is orthogonal $\iff Q^T Q = I$
In other words, $Q^{-1} = Q^T$.

Example

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$ is orthogonal.

Definition

An **orthogonal matrix** is a square matrix Q with orthonormal columns.

It is historical convention to restrict to square matrices, and to say orthogonal matrix even though "orthonormal matrix" might be better.

An $n \times n$ matrix Q is orthogonal $\iff$ $Q^T Q = I$
In other words, $Q^{-1} = Q^T$.

Example

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \text{ is orthogonal.}$$

Why?

Why is $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ orthogonal?

Solution

Why is $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ orthogonal?

Solution

Why is $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ orthogonal?

Solution

Because their columns are a permutation of the standard basis. And so we always have $P^T P = I$.

Why is $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ orthogonal?

Solution

Because their columns are a permutation of the standard basis. And so we always have $P^T P = I$. So what is P^{-1}?

Example

$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal, Why?

Example

$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal, Why?

Solution

- $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ is an orthonormal basis of $\mathbb{R}^2$.

Example

$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal, Why?

Solution

- $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ is an orthonormal basis of $\mathbb{R}^2$.
  Just to make sure: why length 1?

Example

$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal, Why?

Solution

- $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ is an orthonormal basis of $\mathbb{R}^2$.
  Just to make sure: why length 1? Because
  $\| \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \| = \sqrt{\cos^2\theta + \sin^2\theta} = 1.$

Example

$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal, Why?

Solution

- $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ is an orthonormal basis of $\mathbb{R}^2$.
  Just to make sure: why length 1? Because
  $\| \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \| = \sqrt{\cos^2\theta + \sin^2\theta} = 1.$
- Alternatively:
  $$Q^T Q = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Example

$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal, Why?

Solution

- $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ is an orthonormal basis of $\mathbb{R}^2$.
  Just to make sure: why length 1? Because
  $\| \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \| = \sqrt{\cos^2\theta + \sin^2\theta} = 1.$
- Alternatively:
  $$Q^T Q = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

So what is Q^{-1}?

Example

Is $H = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ orthogonal?

Example

Is $H = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ orthogonal?

Solution

No, the columns are orthogonal but not normalized.

Example

Is $H = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ orthogonal?

Solution

No, the columns are orthogonal but not normalized.

But $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is an orthogonal matrix.

Example

(Just for fun) an $n \times n$ matrix with entries ± 1 whose columns are orthogonal is called a *Hadamard matrix* of size n.

Example

(Just for fun) an $n \times n$ matrix with entries ± 1 whose columns are orthogonal is called a *Hadamard matrix* of size n.

Example

(Just for fun) an $n \times n$ matrix with entries ± 1 whose columns are orthogonal is called a *Hadamard matrix* of size n.

A size 4 example: $\begin{bmatrix} H & H \\ H & -H \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Example

(Just for fun) an $n \times n$ matrix with entries ± 1 whose columns are orthogonal is called a *Hadamard matrix* of size n.

A size 4 example: $\begin{bmatrix} H & H \\ H & -H \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Continuing this construction, we get examples of size $8, 16, 32, \ldots$.

It is believed that Hadamard matrices exist for all sizes $4n$.

Example

(Just for fun) an $n \times n$ matrix with entries ± 1 whose columns are orthogonal is called a *Hadamard matrix* of size n.

A size 4 example: $\begin{bmatrix} H & H \\ H & -H \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Continuing this construction, we get examples of size $8, 16, 32, \ldots$.

It is believed that Hadamard matrices exist for all sizes $4n$.

But, no example of size 668 is known yet.

Example

(Just for fun) an $n \times n$ matrix with entries ± 1 whose columns are orthogonal is called a *Hadamard matrix* of size n.

A size 4 example: $\begin{bmatrix} H & H \\ H & -H \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Continuing this construction, we get examples of size $8, 16, 32, \ldots$.
It is believed that Hadamard matrices exist for all sizes $4n$.
But, no example of size 668 is known yet. If you find one you will be famous!