Lecture 27

Math 22 Summer 2017
August 18, 2017

Just for today

- A bit on least-squares
- SVD

§6.5 Theorem 15

§6.5 Theorem 15

Theorem

§6.5 Theorem 15

Theorem

Let A be an $m \times n$ matrix with linearly independent columns.

§6.5 Theorem 15

Theorem

Let A be an $m \times n$ matrix with linearly independent columns. Let $A = QR$ be a QR factorization of A.

§6.5 Theorem 15

Theorem

Let A be an $m \times n$ matrix with linearly independent columns. Let $A = QR$ be a QR factorization of A. Then for every $\mathbf{b} \in \mathbb{R}^m$,

§6.5 Theorem 15

Theorem

Let A be an $m \times n$ matrix with linearly independent columns. Let $A = QR$ be a QR factorization of A. Then for every $\mathbf{b} \in \mathbb{R}^m$, The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution

Theorem

Let A be an $m \times n$ matrix with linearly independent columns. Let $A = QR$ be a QR factorization of A. Then for every $\mathbf{b} \in \mathbb{R}^m$, The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution given by

$$\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}.$$

Theorem

Let A be an $m \times n$ matrix with linearly independent columns. Let $A = QR$ be a QR factorization of A. Then for every $\mathbf{b} \in \mathbb{R}^m$, The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution given by

$$\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}.$$

Proof.

§6.5 Theorem 15

Theorem

Let A be an $m \times n$ matrix with linearly independent columns. Let $A = QR$ be a QR factorization of A. Then for every $\mathbf{b} \in \mathbb{R}^m$, The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution given by

$$\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}.$$

Proof.

By the previous theorem, $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution,

Theorem

Let A be an $m \times n$ matrix with linearly independent columns. Let $A = QR$ be a QR factorization of A. Then for every $\mathbf{b} \in \mathbb{R}^m$, The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution given by

$$\hat{\mathbf{x}} = R^{-1} Q^T \mathbf{b}.$$

Proof.

By the previous theorem, $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution, so we just need to check that the given $\hat{\mathbf{x}}$ works.

Theorem

Let A be an $m \times n$ matrix with linearly independent columns. Let $A = QR$ be a QR factorization of A. Then for every $\mathbf{b} \in \mathbb{R}^m$, The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution given by

$$\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}.$$

Proof.

By the previous theorem, $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution, so we just need to check that the given $\hat{\mathbf{x}}$ works. But

$$\begin{aligned} A\hat{\mathbf{x}} &= QR\hat{\mathbf{x}} \\ &= QRR^{-1}Q^T\mathbf{b} \\ &= QQ^T\mathbf{b} = \hat{\mathbf{b}}. \end{aligned}$$

□

Using the previously computed Q and R,

§6.5 Example ☕ continued

Using the previously computed Q and R, we verify the theorem in this example as follows.

Using the previously computed Q and R, we verify the theorem in this example as follows.

$$\begin{aligned}\hat{\mathbf{x}} &= R^{-1}Q^T\mathbf{b} \\ &= \begin{bmatrix} -1/6 & -1/42 & 1/14 & 5/42 \\ 7/6 & 8/21 & -1/7 & -17/42 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 5/14 \\ 2/7 \end{bmatrix}.\end{aligned}$$

§6.5 Example ☕ continued

Using the previously computed Q and R, we verify the theorem in this example as follows.

$$\begin{aligned}\hat{\mathbf{x}} &= R^{-1}Q^T\mathbf{b} \\ &= \begin{bmatrix} -1/6 & -1/42 & 1/14 & 5/42 \\ 7/6 & 8/21 & -1/7 & -17/42 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 5/14 \\ 2/7 \end{bmatrix}.\end{aligned}$$

Note that this also shows the unilluminating fact that $(A^TA)^{-1}A^T = R^{-1}Q^T$ when defined.

§6.5 Example ☕ concluded

§6.5 Example ☕ concluded

Consider the quantitative data $(2,1), (5,2), (7,3), (8,3)$ in the (t, y)-plane.

Consider the quantitative data $(2,1),(5,2),(7,3),(8,3)$ in the (t,y)-plane. How do we find a line $y = ct + d$ that “best fits” this data?

§6.5 Example ☕ concluded

Consider the quantitative data $(2,1), (5,2), (7,3), (8,3)$ in the (t, y)-plane. How do we find a line $y = ct + d$ that "best fits" this data? Well, to do this we need a notion of error.

§6.5 Example ☕ concluded

Consider the quantitative data $(2, 1), (5, 2), (7, 3), (8, 3)$ in the (t, y)-plane. How do we find a line $y = ct + d$ that “best fits” this data? Well, to do this we need a notion of error. One might measure this error using “squared residuals”.

§6.5 Example ☕ concluded

Consider the quantitative data $(2, 1), (5, 2), (7, 3), (8, 3)$ in the (t, y)-plane. How do we find a line $y = ct + d$ that "best fits" this data? Well, to do this we need a notion of error. One might measure this error using "squared residuals". Draw a picture.

§6.5 Example ☕ concluded

Consider the quantitative data $(2,1),(5,2),(7,3),(8,3)$ in the (t,y)-plane. How do we find a line $y = ct + d$ that "best fits" this data? Well, to do this we need a notion of error. One might measure this error using "squared residuals". Draw a picture. In our case, this error is given by

$$\sum_{i=1}^{4}(y_i - ct_i - d)^2 = (1-2c-d)^2 + (2-5c-d)^2 + (3-7c-d)^2 + (3-8c-3)^2.$$

§6.5 Example concluded

Consider the quantitative data $(2,1),(5,2),(7,3),(8,3)$ in the (t,y)-plane. How do we find a line $y = ct + d$ that "best fits" this data? Well, to do this we need a notion of error. One might measure this error using "squared residuals". Draw a picture. In our case, this error is given by

$$\sum_{i=1}^{4}(y_i - ct_i - d)^2 = (1-2c-d)^2 + (2-5c-d)^2 + (3-7c-d)^2 + (3-8c-3)^2.$$

If we let $A = \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} c \\ d \end{bmatrix},$

§6.5 Example concluded

Consider the quantitative data $(2,1),(5,2),(7,3),(8,3)$ in the (t,y)-plane. How do we find a line $y = ct + d$ that "best fits" this data? Well, to do this we need a notion of error. One might measure this error using "squared residuals". Draw a picture. In our case, this error is given by

$$\sum_{i=1}^{4}(y_i - ct_i - d)^2 = (1-2c-d)^2+(2-5c-d)^2+(3-7c-d)^2+(3-8c-3)^2.$$

If we let $A = \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} c \\ d \end{bmatrix}$, then the above error for a given $\mathbf{x}$ is precisely $\|\mathbf{b} - A\mathbf{x}\|^2$.

§6.5 Example ☕ concluded

Consider the quantitative data $(2,1),(5,2),(7,3),(8,3)$ in the (t,y)-plane. How do we find a line $y = ct + d$ that "best fits" this data? Well, to do this we need a notion of error. One might measure this error using "squared residuals". Draw a picture. In our case, this error is given by

$$\sum_{i=1}^{4}(y_i - ct_i - d)^2 = (1-2c-d)^2 + (2-5c-d)^2 + (3-7c-d)^2 + (3-8c-3)^2.$$

If we let $A = \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} c \\ d \end{bmatrix}$, then the above error for a given $\mathbf{x}$ is precisely $\|\mathbf{b} - A\mathbf{x}\|^2$. This quantity is minimized precisely when

§6.5 Example ☕ concluded

Consider the quantitative data $(2,1),(5,2),(7,3),(8,3)$ in the (t,y)-plane. How do we find a line $y = ct + d$ that "best fits" this data? Well, to do this we need a notion of error. One might measure this error using "squared residuals". Draw a picture. In our case, this error is given by

$$\sum_{i=1}^{4}(y_i - ct_i - d)^2 = (1-2c-d)^2 + (2-5c-d)^2 + (3-7c-d)^2 + (3-8c-3)^2.$$

If we let $A = \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} c \\ d \end{bmatrix}$, then the above error for a given $\mathbf{x}$ is precisely $\|\mathbf{b} - A\mathbf{x}\|^2$. This quantity is minimized precisely when you guessed it

§6.5 Example ☕ concluded

Consider the quantitative data $(2,1),(5,2),(7,3),(8,3)$ in the (t,y)-plane. How do we find a line $y = ct + d$ that "best fits" this data? Well, to do this we need a notion of error. One might measure this error using "squared residuals". Draw a picture. In our case, this error is given by

$$\sum_{i=1}^{4}(y_i - ct_i - d)^2 = (1-2c-d)^2+(2-5c-d)^2+(3-7c-d)^2+(3-8c-3)^2.$$

If we let $A = \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} c \\ d \end{bmatrix}$, then the above error for a given $\mathbf{x}$ is precisely $\|\mathbf{b} - A\mathbf{x}\|^2$. This quantity is minimized precisely when you guessed it $\mathbf{x}$ is a least-squares solution to $A\mathbf{x} = \mathbf{b}$.

§6.5 Example ☕ concluded

Consider the quantitative data $(2,1),(5,2),(7,3),(8,3)$ in the (t,y)-plane. How do we find a line $y = ct + d$ that "best fits" this data? Well, to do this we need a notion of error. One might measure this error using "squared residuals". Draw a picture. In our case, this error is given by

$$\sum_{i=1}^{4}(y_i - ct_i - d)^2 = (1-2c-d)^2 + (2-5c-d)^2 + (3-7c-d)^2 + (3-8c-3)^2.$$

If we let $A = \begin{bmatrix} 2 & 1 \\ 5 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} c \\ d \end{bmatrix}$, then the above error for a given $\mathbf{x}$ is precisely $\|\mathbf{b} - A\mathbf{x}\|^2$. This quantity is minimized precisely when you guessed it $\mathbf{x}$ is a least-squares solution to $A\mathbf{x} = \mathbf{b}$. $\hat{\mathbf{x}} = [5/14 \; 2/7]^T$.

§6.5 Example ☕ concluded

§6.5 Example ☕ concluded

From the least-squares solution we obtain

§6.5 Example ☕ concluded

From the least-squares solution we obtain the line $y = (5/14)t + (2/7)$

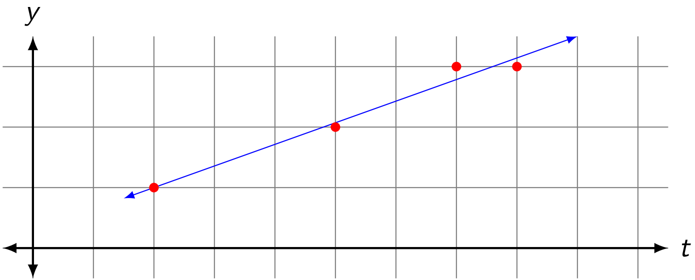
§6.5 Example ☕ concluded

From the least-squares solution we obtain
the line $y = (5/14)t + (2/7)$
that best fits

§6.5 Example ☕ concluded

From the least-squares solution we obtain
the line $y = (5/14)t + (2/7)$
that best fits
the data (2,1),(5,2),(7,3),(8,3) :

From the least-squares solution we obtain
the line $y = (5/14)t + (2/7)$
that best fits
the data (2,1),(5,2),(7,3),(8,3) :



SVD Preliminaries

SVD Preliminaries

Definition

Let A be an $m \times n$ matrix.

SVD Preliminaries

Definition

Let A be an $m \times n$ matrix. We say that A is **symmetric** if $A^T = A$.

SVD Preliminaries

Definition

Let A be an $m \times n$ matrix. We say that A is **symmetric** if $A^T = A$. We say that A is **orthogonally diagonalizable** if $A = PDP^{-1}$ and the columns of P are an orthonormal basis (of eigenvectors).

SVD Preliminaries

Definition

Let A be an $m \times n$ matrix. We say that A is **symmetric** if $A^T = A$. We say that A is **orthogonally diagonalizable** if $A = PDP^{-1}$ and the columns of P are an orthonormal basis (of eigenvectors).

Note that if A is orthogonally diagonalizable, then A is symmetric.

SVD Preliminaries

Definition

Let A be an $m \times n$ matrix. We say that A is **symmetric** if $A^T = A$. We say that A is **orthogonally diagonalizable** if $A = PDP^{-1}$ and the columns of P are an orthonormal basis (of eigenvectors).

Note that if A is orthogonally diagonalizable, then A is symmetric. What's the proof?

SVD Preliminaries

Definition

Let A be an $m \times n$ matrix. We say that A is **symmetric** if $A^T = A$. We say that A is **orthogonally diagonalizable** if $A = PDP^{-1}$ and the columns of P are an orthonormal basis (of eigenvectors).

Note that if A is orthogonally diagonalizable, then A is symmetric. What's the proof? P is an orthogonal matrix.

Definition

Let A be an $m \times n$ matrix. We say that A is **symmetric** if $A^T = A$. We say that A is **orthogonally diagonalizable** if $A = PDP^{-1}$ and the columns of P are an orthonormal basis (of eigenvectors).

Note that if A is orthogonally diagonalizable, then A is symmetric. What's the proof? P is an orthogonal matrix.

Surprisingly the converse is also true!

SVD Preliminaries

Definition

Let A be an $m \times n$ matrix. We say that A is **symmetric** if $A^T = A$. We say that A is **orthogonally diagonalizable** if $A = PDP^{-1}$ and the columns of P are an orthonormal basis (of eigenvectors).

Note that if A is orthogonally diagonalizable, then A is symmetric. What's the proof? P is an orthogonal matrix.

Surprisingly the converse is also true! Emma will speak about this (§7.1 Spectral Theorem) in x-hour on Tuesday.

SVD Preliminaries

Let's see an example of the nontrivial direction.

SVD Preliminaries

Let's see an example of the nontrivial direction.

$$\text{Let } A = \begin{bmatrix} 4 & 3 & 1 & 1 \\ 3 & 4 & 1 & 1 \\ 1 & 1 & 4 & 3 \\ 1 & 1 & 3 & 4 \end{bmatrix}.$$

SVD Preliminaries

Let's see an example of the nontrivial direction.

Let $A = \begin{bmatrix} 4 & 3 & 1 & 1 \\ 3 & 4 & 1 & 1 \\ 1 & 1 & 4 & 3 \\ 1 & 1 & 3 & 4 \end{bmatrix}$. Then

$$A = P \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix} P^T$$

SVD Preliminaries

Let's see an example of the nontrivial direction.

Let $A = \begin{bmatrix} 4 & 3 & 1 & 1 \\ 3 & 4 & 1 & 1 \\ 1 & 1 & 4 & 3 \\ 1 & 1 & 3 & 4 \end{bmatrix}$. Then

$$A = P \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix} P^T$$

with

$$P = \begin{bmatrix} -1/\sqrt{2} & 0 & -1/2 & 1/2 \\ 1/\sqrt{2} & 0 & -1/2 & 1/2 \\ 0 & -1/\sqrt{2} & 1/2 & 1/2 \\ 0 & 1/\sqrt{2} & 1/2 & 1/2 \end{bmatrix}.$$

SVD Preliminaries

SVD Preliminaries

Let A be an $m \times n$ matrix.

SVD Preliminaries

Let A be an $m \times n$ matrix. How do we use results about symmetric matrices when A need not even be square?

SVD Preliminaries

Let A be an $m \times n$ matrix. How do we use results about symmetric matrices when A need not even be square? Let $B = A^T A$.

SVD Preliminaries

Let A be an $m \times n$ matrix. How do we use results about symmetric matrices when A need not even be square? Let $B = A^T A$. Then B is $n \times n$

SVD Preliminaries

Let A be an $m \times n$ matrix. How do we use results about symmetric matrices when A need not even be square? Let $B = A^T A$. Then B is $n \times n$ and

$$B^T = (A^T A)^T = A^T A^{TT} = A^T A = B.$$

SVD Preliminaries

Let A be an $m \times n$ matrix. How do we use results about symmetric matrices when A need not even be square? Let $B = A^T A$. Then B is $n \times n$ and

$$B^T = (A^T A)^T = A^T A^{TT} = A^T A = B.$$

So B is symmetric

SVD Preliminaries

Let A be an $m \times n$ matrix. How do we use results about symmetric matrices when A need not even be square? Let $B = A^T A$. Then B is $n \times n$ and

$$B^T = (A^T A)^T = A^T A^{TT} = A^T A = B.$$

So B is symmetric and therefore orthogonally diagonalizable.

SVD Preliminaries

Let A be an $m \times n$ matrix. How do we use results about symmetric matrices when A need not even be square? Let $B = A^T A$. Then B is $n \times n$ and

$$B^T = (A^T A)^T = A^T A^{TT} = A^T A = B.$$

So B is symmetric and therefore orthogonally diagonalizable.

What does this tell us about A?

SVD Preliminaries

SVD Preliminaries

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B := A^T A$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.

SVD Preliminaries

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B := A^T A$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Again, how do we know such a basis exists?

SVD Preliminaries

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B := A^T A$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Again, how do we know such a basis exists? Spectral Theorem

SVD Preliminaries

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B := A^T A$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Again, how do we know such a basis exists? Spectral Theorem (which we didn't prove...

SVD Preliminaries

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B := A^T A$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Again, how do we know such a basis exists? Spectral Theorem (which we didn't prove... or even state explicitly .

SVD Preliminaries

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B := A^T A$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Again, how do we know such a basis exists? Spectral Theorem (which we didn't prove... or even state explicitly ☹. But Emma will talk about this in the x-hour ☺).

SVD Preliminaries

Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B := A^T A$ with eigenvalues $\lambda_1, \dots, \lambda_n$.

Again, how do we know such a basis exists? Spectral Theorem (which we didn't prove... or even state explicitly ☹. But Emma will talk about this in the x-hour ☺).

Note that for each $\mathbf{v}_i$ we have

$$\begin{aligned} \|A\mathbf{v}_i\|^2 &= (A\mathbf{v}_i) \cdot (A\mathbf{v}_i) \\ &= (A\mathbf{v}_i)^T (A\mathbf{v}_i) \\ &= \mathbf{v}_i^T A^T A \mathbf{v}_i \\ &= \mathbf{v}_i^T B \mathbf{v}_i \\ &= \mathbf{v}_i^T \lambda_i \mathbf{v}_i \\ &= \lambda_i (\mathbf{v}_i \cdot \mathbf{v}_i) = \lambda_i \end{aligned}$$

SVD Preliminaries

Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B := A^T A$ with eigenvalues $\lambda_1, \dots, \lambda_n$.

Again, how do we know such a basis exists? Spectral Theorem (which we didn't prove... or even state explicitly ☹. But Emma will talk about this in the x-hour ☺).

Note that for each $\mathbf{v}_i$ we have

$$\begin{aligned} \|A\mathbf{v}_i\|^2 &= (A\mathbf{v}_i) \cdot (A\mathbf{v}_i) \\ &= (A\mathbf{v}_i)^T (A\mathbf{v}_i) \\ &= \mathbf{v}_i^T A^T A \mathbf{v}_i \\ &= \mathbf{v}_i^T B \mathbf{v}_i \\ &= \mathbf{v}_i^T \lambda_i \mathbf{v}_i \\ &= \lambda_i (\mathbf{v}_i \cdot \mathbf{v}_i) = \lambda_i \end{aligned}$$

Thus every $\lambda_i \geq 0$

SVD Preliminaries

Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $B := A^T A$ with eigenvalues $\lambda_1, \dots, \lambda_n$.

Again, how do we know such a basis exists? Spectral Theorem (which we didn't prove... or even state explicitly ☹. But Emma will talk about this in the x-hour ☺).

Note that for each $\mathbf{v}_i$ we have

$$\begin{aligned}
\|A\mathbf{v}_i\|^2 &= (A\mathbf{v}_i) \cdot (A\mathbf{v}_i) \\
&= (A\mathbf{v}_i)^T (A\mathbf{v}_i) \\
&= \mathbf{v}_i^T A^T A \mathbf{v}_i \\
&= \mathbf{v}_i^T B \mathbf{v}_i \\
&= \mathbf{v}_i^T \lambda_i \mathbf{v}_i \\
&= \lambda_i (\mathbf{v}_i \cdot \mathbf{v}_i) = \lambda_i
\end{aligned}$$

Thus every $\lambda_i \geq 0$ (why?).

SVD Preliminaries

SVD Preliminaries

Thus, we can order the eigenvalues of $A^T A$ as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

SVD Preliminaries

Thus, we can order the eigenvalues of $A^T A$ as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

We define the **singular values** of A to be $\sigma_i := \sqrt{\lambda_i}$

SVD Preliminaries

Thus, we can order the eigenvalues of $A^T A$ as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

We define the **singular values** of A to be $\sigma_i := \sqrt{\lambda_i}$ where $\{\lambda_i\}$ are the eigenvalues of $A^T A$ as above.

SVD Preliminaries

Thus, we can order the eigenvalues of $A^T A$ as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

We define the **singular values** of A to be $\sigma_i := \sqrt{\lambda_i}$ where $\{\lambda_i\}$ are the eigenvalues of $A^T A$ as above. Note that $\sigma_i = \|A\mathbf{v}_i\|$ by our previous computation,

SVD Preliminaries

Thus, we can order the eigenvalues of $A^T A$ as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

We define the **singular values** of A to be $\sigma_i := \sqrt{\lambda_i}$ where $\{\lambda_i\}$ are the eigenvalues of $A^T A$ as above. Note that $\sigma_i = \|A\mathbf{v}_i\|$ by our previous computation, so we have that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

SVD Preliminaries

Thus, we can order the eigenvalues of $A^T A$ as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

We define the **singular values** of A to be $\sigma_i := \sqrt{\lambda_i}$ where $\{\lambda_i\}$ are the eigenvalues of $A^T A$ as above. Note that $\sigma_i = \|A\mathbf{v}_i\|$ by our previous computation, so we have that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

Moreover, we can write

$$\underbrace{\sigma_1, \ldots, \sigma_r}_{>0}, \underbrace{\sigma_{r+1}, \ldots, \sigma_n}_{\text{all zero}}.$$

Thus, we can order the eigenvalues of $A^T A$ as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0.$$

We define the **singular values** of A to be $\sigma_i := \sqrt{\lambda_i}$ where $\{\lambda_i\}$ are the eigenvalues of $A^T A$ as above. Note that $\sigma_i = \|A\mathbf{v}_i\|$ by our previous computation, so we have that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

Moreover, we can write

$$\underbrace{\sigma_1, \ldots, \sigma_r}_{>0}, \underbrace{\sigma_{r+1}, \ldots, \sigma_n}_{\text{all zero}}.$$

The index r above turns out to be the rank of the matrix A.

§7.4 Theorem 9

§7.4 Theorem 9

Theorem

Theorem

Let A be an $m \times n$ matrix.

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$.

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box.

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box. We want to know about the rank of A

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box. We want to know about the rank of A which is the dimension of the column space of A

§7.4 Theorem 9

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box. We want to know about the rank of A which is the dimension of the column space of A which is equal to $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\}$.

§7.4 Theorem 9

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box. We want to know about the rank of A which is the dimension of the column space of A which is equal to $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\}$. But $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\} = \mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$.

§7.4 Theorem 9

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box. We want to know about the rank of A which is the dimension of the column space of A which is equal to $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\}$. But $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\} = \mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. Remember, $\sigma_i = \|A\mathbf{v}_i\|$.

§7.4 Theorem 9

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box. We want to know about the rank of A which is the dimension of the column space of A which is equal to $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\}$. But $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\} = \mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. Remember, $\sigma_i = \|A\mathbf{v}_i\|$. This tells us that the rank of A is $\leq r$.

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box. We want to know about the rank of A which is the dimension of the column space of A which is equal to $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\}$. But $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\} = \mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. Remember, $\sigma_i = \|A\mathbf{v}_i\|$. This tells us that the rank of A is $\leq r$.

It remains to show $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ is linearly independent.

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box. We want to know about the rank of A which is the dimension of the column space of A which is equal to $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\}$. But $\mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\} = \mathrm{Span}\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. Remember, $\sigma_i = \|A\mathbf{v}_i\|$. This tells us that the rank of A is $\leq r$.

It remains to show $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ is linearly independent. To do this we show the vectors are orthogonal:

Theorem

Let A be an $m \times n$ matrix. Then the rank of A is equal to the number of strictly positive singular values of A (i.e. r from the previous slide).

Proof.

We have $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$. Again, this is a black box. We want to know about the rank of A which is the dimension of the column space of A which is equal to $\mathrm{Span}\{A\mathbf{v}_1, \dots, A\mathbf{v}_n\}$. But $\mathrm{Span}\{A\mathbf{v}_1, \dots, A\mathbf{v}_n\} = \mathrm{Span}\{A\mathbf{v}_1, \dots, A\mathbf{v}_r\}$. Remember, $\sigma_i = \|A\mathbf{v}_i\|$. This tells us that the rank of A is $\leq r$.

It remains to show $\{A\mathbf{v}_1, \dots, A\mathbf{v}_r\}$ is linearly independent. To do this we show the vectors are orthogonal: $(A\mathbf{v}_i) \cdot (A\mathbf{v}_j) = (A\mathbf{v}_i)^T (A\mathbf{v}_j) = \mathbf{v}_i^T A^T A \mathbf{v}_j = \mathbf{v}_i^T \lambda_j \mathbf{v}_j = \lambda_j (\mathbf{v}_i \cdot \mathbf{v}_j) = 0$. □

§7.4 Singular value decomposition

Definition

§7.4 Singular value decomposition

Definition

The **singular value decomposition** (SVD) of an $m \times n$ matrix A is

$$A = U\Sigma V^T$$

§7.4 Singular value decomposition

Definition

The **singular value decomposition** (SVD) of an $m \times n$ matrix A is

$$A = U\Sigma V^T$$

where U is an $m \times m$ orthogonal matrix,

Definition

The **singular value decomposition** (SVD) of an $m \times n$ matrix A is

$$A = U\Sigma V^T$$

where U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix whose columns consist of eigenvectors of $A^T A$,

Definition

The **singular value decomposition** (SVD) of an $m \times n$ matrix A is

$$A = U\Sigma V^T$$

where U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix whose columns consist of eigenvectors of A^TA, and Σ is an $m \times n$ "diagonalish" matrix:

Definition

The **singular value decomposition** (SVD) of an $m \times n$ matrix A is

$$A = U \Sigma V^T$$

where U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix whose columns consist of eigenvectors of $A^T A$, and Σ is an $m \times n$ "diagonalish" matrix: $\Sigma = \begin{bmatrix} D & 0 \\ 0 & \end{bmatrix}$

§7.4 Singular value decomposition

Definition

The **singular value decomposition** (SVD) of an $m \times n$ matrix A is

$$A = U\Sigma V^T$$

where U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix whose columns consist of eigenvectors of A^TA, and Σ is an $m \times n$ “diagonalish” matrix: $\Sigma = \begin{bmatrix} D & 0 \\ 0 & \end{bmatrix}$

where D is the $r \times r$ matrix

$$D = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}$$

Definition

The **singular value decomposition** (SVD) of an $m \times n$ matrix A is

$$A = U\Sigma V^T$$

where U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix whose columns consist of eigenvectors of A^TA, and Σ is an $m \times n$ "diagonalish" matrix: $\Sigma = \begin{bmatrix} D & 0 \\ 0 & \end{bmatrix}$ 

where D is the $r \times r$ matrix

$$D = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}$$

and  is the $(m-r) \times (n-r)$ matrix of zeros.

§7.4 Singular value decomposition

Note that U and V are orthogonal matrices.

Note that U and V are orthogonal matrices. The columns of U are called the **left singular vectors** of A.

Note that U and V are orthogonal matrices. The columns of U are called the **left singular vectors** of A. The columns of V are called the **right singular vectors** of A.

Note that U and V are orthogonal matrices. The columns of U are called the **left singular vectors** of A. The columns of V are called the **right singular vectors** of A.

Also note that the U and V are not uniquely determined,

Note that U and V are orthogonal matrices. The columns of U are called the **left singular vectors** of A. The columns of V are called the **right singular vectors** of A.

Also note that the U and V are not uniquely determined, but the singular values are unique.

Note that U and V are orthogonal matrices. The columns of U are called the **left singular vectors** of A. The columns of V are called the **right singular vectors** of A.

Also note that the U and V are not uniquely determined, but the singular values are unique.

We have already discussed how to find V.

Note that U and V are orthogonal matrices. The columns of U are called the **left singular vectors** of A. The columns of V are called the **right singular vectors** of A.

Also note that the U and V are not uniquely determined, but the singular values are unique.

We have already discussed how to find V. Next we will show how we get U,

Note that U and V are orthogonal matrices. The columns of U are called the **left singular vectors** of A. The columns of V are called the **right singular vectors** of A.

Also note that the U and V are not uniquely determined, but the singular values are unique.

We have already discussed how to find V. Next we will show how we get U, and then we will prove that the story checks out!

§7.4 Theorem 10

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$.

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?
It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

§7.4 Theorem 10

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?
It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?
It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col}A$.

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?
It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col}A$. But U is an $m \times m$ matrix and we might have $m > r$.

§7.4 Theorem 10

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set? It is an orthogonal basis for $\mathrm{Col} A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col} A$. But U is an $m \times m$ matrix and we might have $m > r$. So what do we do?

§7.4 Theorem 10

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?
It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col}A$. But U is an $m \times m$ matrix and we might have $m > r$. So what do we do?
Extend the basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ of $\mathrm{Col}A$ to a basis of $\mathbb{R}^m$.

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set? It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col}A$. But U is an $m \times m$ matrix and we might have $m > r$. So what do we do? Extend the basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ of $\mathrm{Col}A$ to a basis of $\mathbb{R}^m$. Call it $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$.

§7.4 Theorem 10

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?
It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col}A$. But U is an $m \times m$ matrix and we might have $m > r$. So what do we do? Extend the basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ of $\mathrm{Col}A$ to a basis of $\mathbb{R}^m$. Call it $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$. Then finally we get $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m]$.

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?
It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col}A$. But U is an $m \times m$ matrix and we might have $m > r$. So what do we do? Extend the basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ of $\mathrm{Col}A$ to a basis of $\mathbb{R}^m$. Call it $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$. Then finally we get $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m]$.

Theorem

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?
It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col}A$. But U is an $m \times m$ matrix and we might have $m > r$. So what do we do? Extend the basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ of $\mathrm{Col}A$ to a basis of $\mathbb{R}^m$. Call it $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$. Then finally we get $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m]$.

Theorem

Let A be an $m \times n$ matrix with rank r.

§7.4 Theorem 10

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set?
It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col}A$. But U is an $m \times m$ matrix and we might have $m > r$. So what do we do? Extend the basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ of $\mathrm{Col}A$ to a basis of $\mathbb{R}^m$. Call it $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$. Then finally we get $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m]$.

Theorem

Let A be an $m \times n$ matrix with rank r. Then $A = U\Sigma V^T$ (with U, Σ, V defined above) is a singular value decomposition of A.

Consider $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$. What do we know about this set? It is an orthogonal basis for $\mathrm{Col}A \subseteq \mathbb{R}^m$.

Let

$$\mathbf{u}_i := \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is an orthonormal basis for $\mathrm{Col}A$. But U is an $m \times m$ matrix and we might have $m > r$. So what do we do? Extend the basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ of $\mathrm{Col}A$ to a basis of $\mathbb{R}^m$. Call it $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$. Then finally we get $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m]$.

Theorem

Let A be an $m \times n$ matrix with rank r. Then $A = U\Sigma V^T$ (with U, Σ, V defined above) is a singular value decomposition of A.

Let's prove that this all checks out.

§7.4 Proof of Theorem 10

§7.4 Proof of Theorem 10

First we compute AV:

§7.4 Proof of Theorem 10

First we compute AV:

$$\begin{aligned} AV &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_n] \\ &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}] \\ &= [\sigma_1\mathbf{u}_1 \ \cdots \ \sigma_r\mathbf{u}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}] \end{aligned}$$

§7.4 Proof of Theorem 10

First we compute AV:

$$\begin{aligned} AV &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_n] \\ &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}] \\ &= [\sigma_1\mathbf{u}_1 \ \cdots \ \sigma_r\mathbf{u}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}] \end{aligned}$$

Now compute $U\Sigma$:

§7.4 Proof of Theorem 10

First we compute AV:

$$
\begin{aligned}
AV &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_n] \\
&= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}] \\
&= [\sigma_1\mathbf{u}_1 \ \cdots \ \sigma_r\mathbf{u}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}]
\end{aligned}
$$

Now compute $U\Sigma$:

$$
U\Sigma = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m] \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & \end{array}\right] = [\sigma_1\mathbf{u}_1 \ \cdots \ \sigma_r\mathbf{u}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}]
$$

§7.4 Proof of Theorem 10

First we compute AV:

$$\begin{aligned} AV &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_n] \\ &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}] \\ &= [\sigma_1\mathbf{u}_1 \ \cdots \ \sigma_r\mathbf{u}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}] \end{aligned}$$

Now compute $U\Sigma$:

$$U\Sigma = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m] \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right] = [\sigma_1\mathbf{u}_1 \ \cdots \ \sigma_r\mathbf{u}_r \ \underbrace{\mathbf{0} \ \cdots \ \mathbf{0}}_{n-r}]$$

Thus

$$U\Sigma V^T = (U\Sigma)V^T = (AV)V^T = A(VV^T) = AI_n = A.$$

§7.4 SVD Example

§7.4 SVD Example ☃

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$

§7.4 SVD Example ☃

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$

§7.4 SVD Example

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^TA = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ Looks symmetric.

§7.4 SVD Example

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ Looks symmetric. The charpoly of $A^T A$ is $(\lambda - 90)(\lambda - 10)$, so what are the singular values?

§7.4 SVD Example

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ Looks symmetric. The charpoly of $A^T A$ is $(\lambda - 90)(\lambda - 10)$, so what are the singular values? $\sigma_1 = \sqrt{90}$, $\sigma_2 = \sqrt{10}$.

§7.4 SVD Example

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ Looks symmetric. The charpoly of $A^T A$ is $(\lambda - 90)(\lambda - 10)$, so what are the singular values? $\sigma_1 = \sqrt{90}$, $\sigma_2 = \sqrt{10}$. Thus

$$\Sigma = \begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}.$$

§7.4 SVD Example

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ Looks symmetric. The charpoly of $A^T A$ is $(\lambda - 90)(\lambda - 10)$, so what are the singular values? $\sigma_1 = \sqrt{90}$, $\sigma_2 = \sqrt{10}$. Thus

$$\Sigma = \begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}.$$

How do we find V?

§7.4 SVD Example

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ Looks symmetric. The charpoly of $A^T A$ is $(\lambda - 90)(\lambda - 10)$, so what are the singular values? $\sigma_1 = \sqrt{90}$, $\sigma_2 = \sqrt{10}$. Thus

$$\Sigma = \begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}.$$

How do we find V? The columns of V are normalized eigenvectors of $A^T A$.

§7.4 SVD Example

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^TA = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ Looks symmetric. The charpoly of A^TA is $(\lambda - 90)(\lambda - 10)$, so what are the singular values? $\sigma_1 = \sqrt{90}$, $\sigma_2 = \sqrt{10}$. Thus

$$\Sigma = \begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}.$$

How do we find V? The columns of V are normalized eigenvectors of A^TA. A basis for $\mathrm{Nul}(A^TA - 90I_2)$ is $\{[2\ 1]^T\}$.

§7.4 SVD Example

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ Looks symmetric. The charpoly of $A^T A$ is $(\lambda - 90)(\lambda - 10)$, so what are the singular values? $\sigma_1 = \sqrt{90}$, $\sigma_2 = \sqrt{10}$. Thus

$$\Sigma = \begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}.$$

How do we find V? The columns of V are normalized eigenvectors of $A^T A$. A basis for $\mathrm{Nul}(A^T A - 90 I_2)$ is $\{[2\ 1]^T\}$. A basis for $\mathrm{Nul}(A^T A - 10 I_2)$ is $\{[-1\ 2]^T\}$.

§7.4 SVD Example

Let $A = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$ Then $A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$ Looks symmetric. The charpoly of $A^T A$ is $(\lambda - 90)(\lambda - 10)$, so what are the singular values? $\sigma_1 = \sqrt{90}$, $\sigma_2 = \sqrt{10}$. Thus

$$\Sigma = \begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}.$$

How do we find V? The columns of V are normalized eigenvectors of $A^T A$. A basis for $\mathrm{Nul}(A^T A - 90 I_2)$ is $\{[2\ 1]^T\}$. A basis for $\mathrm{Nul}(A^T A - 10 I_2)$ is $\{[-1\ 2]^T\}$. Normalizing we get that

$$V = [\mathbf{v}_1\ \mathbf{v}_2] = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}.$$

§7.4 Example ☃ continued

§7.4 Example ☃ continued

We have $\Sigma_{3\times 2}$ and $V_{2\times 2}$.

§7.4 Example ☃ continued

We have $\Sigma_{3\times 2}$ and $V_{2\times 2}$. We need to compute $U_{3\times 3}$.

§7.4 Example ☃ continued

We have $\Sigma_{3\times 2}$ and $V_{2\times 2}$. We need to compute $U_{3\times 3}$. To do this we compute:

$$A\mathbf{v}_1 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5} \\ 3\sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_1 = \frac{A\mathbf{v}_1}{\|A\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$A\mathbf{v}_2 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -\sqrt{5} \\ \sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\|A\mathbf{v}_2\|} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

§7.4 Example ☃ continued

We have $\Sigma_{3\times 2}$ and $V_{2\times 2}$. We need to compute $U_{3\times 3}$. To do this we compute:

$$A\mathbf{v}_1 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5} \\ 3\sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_1 = \frac{A\mathbf{v}_1}{\|A\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$A\mathbf{v}_2 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -\sqrt{5} \\ \sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\|A\mathbf{v}_2\|} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

Is $\{\mathbf{u}_1, \mathbf{u}_2\}$ a basis for $\mathbb{R}^3$?

§7.4 Example ☃ continued

We have $\Sigma_{3\times 2}$ and $V_{2\times 2}$. We need to compute $U_{3\times 3}$. To do this we compute:

$$A\mathbf{v}_1 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5} \\ 3\sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_1 = \frac{A\mathbf{v}_1}{\|A\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$A\mathbf{v}_2 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -\sqrt{5} \\ \sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\|A\mathbf{v}_2\|} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

Is $\{\mathbf{u}_1, \mathbf{u}_2\}$ a basis for $\mathbb{R}^3$? No!

We have $\Sigma_{3\times 2}$ and $V_{2\times 2}$. We need to compute $U_{3\times 3}$. To do this we compute:

$$A\mathbf{v}_1 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5} \\ 3\sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_1 = \frac{A\mathbf{v}_1}{\|A\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$A\mathbf{v}_2 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -\sqrt{5} \\ \sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\|A\mathbf{v}_2\|} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

Is $\{\mathbf{u}_1, \mathbf{u}_2\}$ a basis for $\mathbb{R}^3$? No! We need to extend it to a basis of $\mathbb{R}^3$.

We have $\Sigma_{3\times 2}$ and $V_{2\times 2}$. We need to compute $U_{3\times 3}$. To do this we compute:

$$A\mathbf{v}_1 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5} \\ 3\sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_1 = \frac{A\mathbf{v}_1}{\|A\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$A\mathbf{v}_2 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -\sqrt{5} \\ \sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\|A\mathbf{v}_2\|} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

Is $\{\mathbf{u}_1, \mathbf{u}_2\}$ a basis for $\mathbb{R}^3$? No! We need to extend it to a basis of $\mathbb{R}^3$. So what is $\mathbf{u}_3$?

We have $\Sigma_{3\times 2}$ and $V_{2\times 2}$. We need to compute $U_{3\times 3}$. To do this we compute:

$$A\mathbf{v}_1 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5} \\ 3\sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_1 = \frac{A\mathbf{v}_1}{\|A\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$A\mathbf{v}_2 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -\sqrt{5} \\ \sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\|A\mathbf{v}_2\|} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

Is $\{\mathbf{u}_1, \mathbf{u}_2\}$ a basis for $\mathbb{R}^3$? No! We need to extend it to a basis of $\mathbb{R}^3$. So what is $\mathbf{u}_3$? $\mathbf{u}_3 = [0\ 0\ 1]^T$.

§7.4 Example ☃ continued

We have $\Sigma_{3\times 2}$ and $V_{2\times 2}$. We need to compute $U_{3\times 3}$. To do this we compute:

$$A\mathbf{v}_1 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5} \\ 3\sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_1 = \frac{A\mathbf{v}_1}{\|A\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$A\mathbf{v}_2 = \begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -\sqrt{5} \\ \sqrt{5} \\ 0 \end{bmatrix} \implies \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\|A\mathbf{v}_2\|} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

Is $\{\mathbf{u}_1, \mathbf{u}_2\}$ a basis for $\mathbb{R}^3$? No! We need to extend it to a basis of $\mathbb{R}^3$. So what is $\mathbf{u}_3$? $\mathbf{u}_3 = [0\ 0\ 1]^T$. Then $U = [\mathbf{u}_1\ \mathbf{u}_2\ \mathbf{u}_3]$.

§7.4 Example ☃ concluded

§7.4 Example ☃ concluded

We can now verify that our singular value decomposition of A works:

We can now verify that our singular value decomposition of A works:

$$\underbrace{\begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}}_{A_{3\times 2}} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{U_{3\times 3}} \underbrace{\begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}}_{\Sigma_{3\times 2}} \underbrace{\begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}^T}_{V^T_{2\times 2}}$$

We can now verify that our singular value decomposition of A works:

$$\underbrace{\begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}}_{A_{3\times 2}} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{U_{3\times 3}} \underbrace{\begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}}_{\Sigma_{3\times 2}} \underbrace{\begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}^T}_{V^T_{2\times 2}}$$

We will conclude this course by explaining some of the applications of SVD. So stay tuned.

§7.4 Example ☃ concluded

We can now verify that our singular value decomposition of A works:

$$\underbrace{\begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}}_{A_{3\times 2}} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{U_{3\times 3}} \underbrace{\begin{bmatrix} \sqrt{90} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}}_{\Sigma_{3\times 2}} \underbrace{\begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}^T}_{V^T_{2\times 2}}$$

We will conclude this course by explaining some of the applications of SVD. So stay tuned.

For some practice with small 2×2 examples you might want to take a look at `https://goo.gl/oiFXd8`