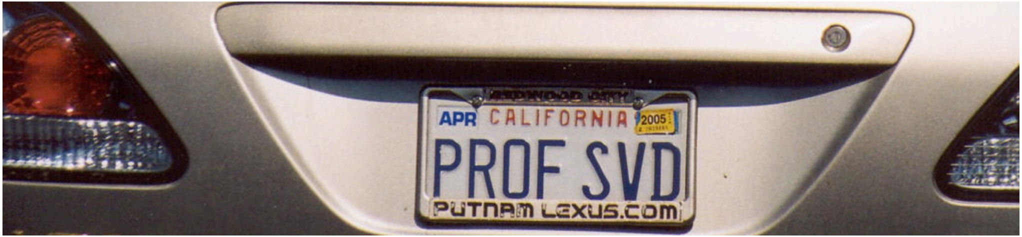
Quiz

1. Write a procedure `check_least_squares`$(A, \hat{\mathbf{x}}, \mathbf{b})$ with the following spec:
   - *input:* Mat A, Vec $\mathbf{u}$, Vec $\mathbf{b}$
   - *output:* `True` if $\mathbf{u}$ is the solution to the least-squares problem $A\mathbf{x} \approx \mathbf{b}$, i.e. if $\mathbf{u}$ minimizes $\|\mathbf{b} - A\mathbf{u}\|^2$.

   Assume that the vectors are legal, i.e. the domain of $\mathbf{u}$ equals the column label set of A and the domain of $\mathbf{b}$ equals the row label set of A. Also assume that there is no floating-point error, i.e. that all calculations are precisely correct. Do not assume that the columns of A are linearly independent. Your procedure should not explicitly use any other procedures. (Of course, it can use the usual operations on matrices and vectors.)
2. Suppose $\mathcal{U}$ and $\mathcal{V}$ are subspaces of $\mathcal{W}$. What does it mean to say that $\mathcal{V}$ is the orthogonal complement of $\mathcal{U}$ in $\mathcal{W}$? Give the definition.

[11] The Singular Value Decomposition

The Singular Value Decomposition



Gene Golub's license plate, photographed by Professor P. M. Kroonenberg of Leiden University.

Frobenius norm for matrices

We have defined a norm for vectors over $\mathbb{R}$: $\|[x_1, x_2, \ldots, x_n]\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

Now we define a norm for matrices: interpret the matrix as a vector.

$$\|A\|_F = \sqrt{\text{sum of squares of elements of } A}$$

called the *Frobenius norm* of a matrix.

Squared norm is just sum of squares of the elements.

Example: $\left\| \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \right\|_F^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2$

Can group in terms of rows .... or columns

$$\left\| \left[\begin{array}{ccc} 1 & 2 & 3 \\ \hline 4 & 5 & 6 \end{array}\right] \right\|_F^2 = (1^2 + 2^2 + 3^2) + (4^2 + 5^2 + 6^2) = \|[1, 2, 3]\|^2 + \|[4, 5, 6]\|^2$$

$$\left\| \left[\begin{array}{c|c|c} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}\right] \right\|_F^2 = (1^2 + 4^2) + (2^2 + 5^2) + (3^2 + 6^2) = \|[1, 4]\|^2 + \|[2, 5]\|^2 + \|[3, 6]\|^2$$

Frobenius norm for matrices

Example: $\left\| \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \right\|_F^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2$

Can group in terms of rows .... or columns

$$\left\| \left[\begin{array}{ccc} 1 & 2 & 3 \\ \hline 4 & 5 & 6 \end{array}\right] \right\|_F^2 = (1^2 + 2^2 + 3^2) + (4^2 + 5^2 + 6^2) = \|[1,2,3]\|^2 + \|[4,5,6]\|^2$$

$$\left\| \left[\begin{array}{c|c|c} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}\right] \right\|_F^2 = (1^2 + 4^2) + (2^2 + 5^2) + (3^2 + 6^2) = \|[1,4]\|^2 + \|[2,5]\|^2 + \|[3,6]\|^2$$

Proposition: Squared Frobenius norm of a matrix is the sum of the squared norms of its rows ...

$$\left\| \left[\begin{array}{c} \mathbf{a}_1 \\ \hline \vdots \\ \hline \mathbf{a}_m \end{array}\right] \right\|_F^2 = \|\mathbf{a}_1\|^2 + \cdots + \|\mathbf{a}_m\|^2$$

Frobenius norm for matrices

Example: $\left\| \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \right\|_F^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2$

Can group in terms of rows .... or columns

$$\left\| \begin{bmatrix} 1 & 2 & 3 \\ \hline 4 & 5 & 6 \end{bmatrix} \right\|_F^2 = (1^2 + 2^2 + 3^2) + (4^2 + 5^2 + 6^2) = \|[1,2,3]\|^2 + \|[4,5,6]\|^2$$

$$\left\| \left[\begin{array}{c|c|c} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array} \right] \right\|_F^2 = (1^2 + 4^2) + (2^2 + 5^2) + (3^2 + 6^2) = \|[1,4]\|^2 + \|[2,5]\|^2 + \|[3,6]\|^2$$

Proposition: Squared Frobenius norm of a matrix is the sum of the squared norms of its rows ... or of its columns.

$$\left\| \left[\begin{array}{c|c|c} & & \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ & & \end{array} \right] \right\|_F^2 = \|\mathbf{v}_1\|^2 + \cdots + \|\mathbf{v}_n\|^2$$

Low-rank matrices

Saving space and saving time

$$\begin{bmatrix} \\ \mathbf{u} \\ \\ \end{bmatrix} \begin{bmatrix} & \mathbf{v}^T & \end{bmatrix}$$

$$\left(\begin{bmatrix} \\ \mathbf{u} \\ \\ \end{bmatrix} \begin{bmatrix} & \mathbf{v}^T & \end{bmatrix} \right) \begin{bmatrix} \\ \mathbf{w} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \mathbf{u} \\ \\ \end{bmatrix} \left(\begin{bmatrix} & \mathbf{v}^T & \end{bmatrix} \begin{bmatrix} \\ \mathbf{w} \\ \\ \end{bmatrix} \right)$$

$$\left[\begin{array}{c|c} & \\ \mathbf{u}_1 & \mathbf{u}_2 \\ & \end{array} \right] \left[\begin{array}{c} \mathbf{v}_1^T \\ \hline \mathbf{v}_2^T \end{array} \right]$$

Silly compression

Represent a grayscale $m \times n$ image by an $m \times n$ matrix A. (Requires mn numbers to represent.)
Find a low-rank matrix $\tilde{A}$ that is as close as possible to A. (For rank r, requires only $r(m+n)$ numbers to represent.)
Original image (625×1024, so about 625k numbers)

Silly compression

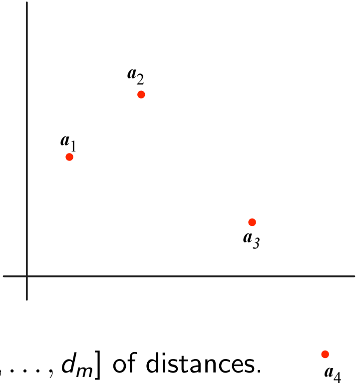
Represent a grayscale $m \times n$ image by an $m \times n$ matrix A. (Requires mn numbers to represent.)
Find a low-rank matrix $\tilde{A}$ that is as close as possible to A. (For rank r, requires only $r(m+n)$ numbers to represent.)
Rank-50 approximation (so about 82k numbers)

The *trolley-line-location* problem

Given the locations of m houses $\mathbf{a}_1, \ldots, \mathbf{a}_m$,
we must choose where to run a trolley line.

The trolley line must go through downtown (origin)
and must be a straight line.

The goal is to locate the trolley line so that it is as
close as possible to the m houses.



Specify line by unit-norm vector $\mathbf{v}$: line is Span $\{\mathbf{v}\}$.

In measuring objective, how to combine individual objectives?

As in least squares, we minimize the 2-norm of the vector $[d_1, \ldots, d_m]$ of distances.

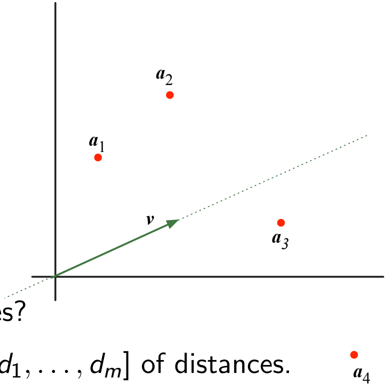
Equivalent to minimizing the square of the 2-norm of this vector, i.e. $d_1^2 + \cdots + d_m^2$.

The *trolley-line-location* problem

Given the locations of m houses $\mathbf{a}_1, \ldots, \mathbf{a}_m$,
we must choose where to run a trolley line.

The trolley line must go through downtown (origin)
and must be a straight line.

The goal is to locate the trolley line so that it is as
close as possible to the m houses.



Specify line by unit-norm vector $\mathbf{v}$: line is Span $\{\mathbf{v}\}$.

In measuring objective, how to combine individual objectives?

As in least squares, we minimize the 2-norm of the vector $[d_1, \ldots, d_m]$ of distances.

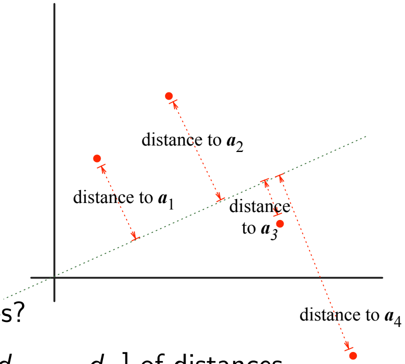
Equivalent to minimizing the square of the 2-norm of this vector, i.e. $d_1^2 + \cdots + d_m^2$.

The *trolley-line-location* problem

Given the locations of m houses $\mathbf{a}_1, \ldots, \mathbf{a}_m$, we must choose where to run a trolley line.

The trolley line must go through downtown (origin) and must be a straight line.

The goal is to locate the trolley line so that it is as close as possible to the m houses.



Specify line by unit-norm vector $\mathbf{v}$: line is Span $\{\mathbf{v}\}$.

In measuring objective, how to combine individual objectives?

As in least squares, we minimize the 2-norm of the vector $[d_1, \ldots, d_m]$ of distances.

Equivalent to minimizing the square of the 2-norm of this vector, i.e. $d_1^2 + \cdots + d_m^2$.

Solution to the *trolley-line-location* problem

For each vector $\mathbf{a}_i$, write $\mathbf{a}_i = \mathbf{a}_i^{\|\mathbf{v}} + \mathbf{a}_i^{\perp\mathbf{v}}$ where $\mathbf{a}_i^{\|\mathbf{v}}$ is the projection of $\mathbf{a}_i$ along $\mathbf{v}$ and $\mathbf{a}_i^{\perp\mathbf{v}}$ is the projection orthogonal to $\mathbf{v}$.

$$\begin{aligned} \mathbf{a}_1^{\perp\mathbf{v}} &= \mathbf{a}_1 - \mathbf{a}_1^{\|\mathbf{v}} \\ &\vdots \\ \mathbf{a}_m^{\perp\mathbf{v}} &= \mathbf{a}_m - \mathbf{a}_m^{\|\mathbf{v}} \end{aligned}$$

By the Pythagorean Theorem,

$$\begin{aligned} \|\mathbf{a}_1^{\perp\mathbf{v}}\|^2 &= \|\mathbf{a}_1\|^2 - \|\mathbf{a}_1^{\|\mathbf{v}}\|^2 \\ &\vdots \\ \|\mathbf{a}_m^{\perp\mathbf{v}}\|^2 &= \|\mathbf{a}_m\|^2 - \|\mathbf{a}_m^{\|\mathbf{v}}\|^2 \end{aligned}$$

Since the distance from $\mathbf{a}_i$ to Span $\{\mathbf{v}\}$ is $\|\mathbf{a}_i^{\perp\mathbf{v}}\|$, we have

$$\begin{aligned} (\text{dist from } \mathbf{a}_1 \text{ to Span } \{\mathbf{v}\})^2 &= \|\mathbf{a}_1\|^2 - \|\mathbf{a}_1^{\|\mathbf{v}}\|^2 \\ &\vdots \\ (\text{dist from } \mathbf{a}_m \text{ to Span } \{\mathbf{v}\})^2 &= \|\mathbf{a}_m\|^2 - \|\mathbf{a}_m^{\|\mathbf{v}}\|^2 \end{aligned}$$

Solution to the *trolley-line-location* problem

$$\mathbf{a}_1^{\perp \mathbf{v}} = \mathbf{a}_1 - \mathbf{a}_1^{\| \mathbf{v}}$$
$$\vdots$$
$$\mathbf{a}_m^{\perp \mathbf{v}} = \mathbf{a}_m - \mathbf{a}_m^{\| \mathbf{v}}$$

By the Pythagorean Theorem,

$$\begin{aligned} \|\mathbf{a}_1^{\perp \mathbf{v}}\|^2 &= \|\mathbf{a}_1\|^2 - \|\mathbf{a}_1^{\| \mathbf{v}}\|^2 \\ &\vdots \\ \|\mathbf{a}_m^{\perp \mathbf{v}}\|^2 &= \|\mathbf{a}_m\|^2 - \|\mathbf{a}_m^{\| \mathbf{v}}\|^2 \end{aligned}$$

Since the distance from $\mathbf{a}_i$ to Span $\{\mathbf{v}\}$ is $\|\mathbf{a}_i^{\perp \mathbf{v}}\|$, we have

$$\begin{array}{rcccc} (\text{dist from } \mathbf{a}_1 \text{ to Span } \{\mathbf{v}\})^2 & = & \|\mathbf{a}_1\|^2 & - & \|\mathbf{a}_1^{\| \mathbf{v}}\|^2 \\ & \vdots & & & \\ (\text{dist from } \mathbf{a}_m \text{ to Span } \{\mathbf{v}\})^2 & = & \|\mathbf{a}_m\|^2 & - & \|\mathbf{a}_m^{\| \mathbf{v}}\|^2 \\ \hline \sum_i (\text{dist from } \mathbf{a}_i \text{ to Span } \{\mathbf{v}\})^2 & = & \|\mathbf{a}_1\|^2 + \cdots + \|\mathbf{a}_m\|^2 & - & \left(\|\mathbf{a}_1^{\| \mathbf{v}}\|^2 + \cdots + \|\mathbf{a}_m^{\| \mathbf{v}}\|^2\right) \\ & = & \|A\|_F^2 & - & \left(\langle \mathbf{a}_1, \mathbf{v} \rangle^2 + \cdots + \langle \mathbf{a}_m, \mathbf{v} \rangle^2 \right) \end{array}$$

using $\mathbf{a}_i^{\| \mathbf{v}} = \langle \mathbf{a}_i, \mathbf{v} \rangle \ \mathbf{v}$ and hence $\|\mathbf{a}_i^{\| \mathbf{v}}\|^2 = \langle \mathbf{a}_i, \mathbf{v} \rangle^2 \ \|\mathbf{v}\|^2 = \langle \mathbf{a}_i, \mathbf{v} \rangle^2$

Solution to the *trolley-line-location* problem, continued

By dot-product interpretation of matrix-vector multiplication,

$$\begin{bmatrix} \mathbf{a}_1 \\ \hline \vdots \\ \hline \mathbf{a}_m \end{bmatrix} \begin{bmatrix} \\ \mathbf{v} \\ \\ \end{bmatrix} = \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{a}_m, \mathbf{v} \rangle \end{bmatrix} \tag{1}$$

so

$$\|A\mathbf{v}\|^2 = \left(\langle \mathbf{a}_1, \mathbf{v} \rangle^2 + \langle \mathbf{a}_2, \mathbf{v} \rangle^2 + \cdots + \langle \mathbf{a}_m, \mathbf{v} \rangle^2 \right)$$

We get

$$\sum_i (\text{distance from } \mathbf{a}_i \text{ to Span } \{\mathbf{v}\})^2 \;=\; ||A||_F^2 \;-\; \|A\mathbf{v}\|^2$$

Therefore best vector $\mathbf{v}$ is a unit vector that maximizes $||A\mathbf{v}||^2$ (equiv., maximizes $||A\mathbf{v}||$).

Solution to the *trolley-line-location* problem, continued

$$\sum_i (\text{dist from } \mathbf{a}_i \text{ to Span } \{\mathbf{v}\})^2 = \|A\|_F^2 - \left(\langle \mathbf{a}_1, \mathbf{v} \rangle^2 + \cdots + \langle \mathbf{a}_m, \mathbf{v} \rangle^2 \right)$$

By dot-product interpretation of matrix-vector multiplication,

$$\begin{bmatrix} \mathbf{a}_1 \\ \hline \vdots \\ \hline \mathbf{a}_m \end{bmatrix} \begin{bmatrix} \\ \mathbf{v} \\ \\ \end{bmatrix} = \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{v} \rangle \\ \vdots \\ \langle \mathbf{a}_m, \mathbf{v} \rangle \end{bmatrix} \tag{1}$$

so

$$\|A\mathbf{v}\|^2 = \left(\langle \mathbf{a}_1, \mathbf{v} \rangle^2 + \langle \mathbf{a}_2, \mathbf{v} \rangle^2 + \cdots + \langle \mathbf{a}_m, \mathbf{v} \rangle^2 \right)$$

We get

$$\sum_i (\text{distance from } \mathbf{a}_i \text{ to Span } \{\mathbf{v}\})^2 \;=\; ||A||_F^2 \;-\; ||A\mathbf{v}||^2$$

Therefore best vector $\mathbf{v}$ is a unit vector that maximizes $||A\mathbf{v}||^2$ (equiv., maximizes $||A\mathbf{v}||$).

Solution to the *trolley-line-location* problem, continued

$$\sum_i (\text{distance from } \mathbf{a}_i \text{ to Span } \{\mathbf{v}\})^2 \quad = \quad ||A||_F^2 \quad - \quad \|A\mathbf{v}\|^2$$

Therefore best vector $\mathbf{v}$ is a unit vector that maximizes $||A\mathbf{v}||^2$ (equiv., maximizes $||A\mathbf{v}||$).

def trolley_line_location(A):
$\mathbf{v}_1 = \arg\max\{||A\mathbf{v}|| \; : \; ||\mathbf{v}|| = 1\}$
$\sigma_1 = ||A\mathbf{v}_1||$
return $\mathbf{v}_1$

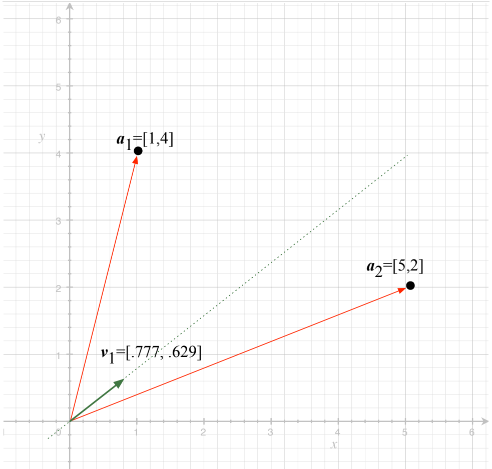
So far, this is a solution only in *principle* since we have not specified how to actually compute $\mathbf{v}_1$.

Definition: σ_1 is *first singular value* of A, and $\mathbf{v}_1$ is *first right singular vector*.

Trolley-line-location problem, example

Example: Let $A = \begin{bmatrix} 1 & 4 \\ 5 & 2 \end{bmatrix}$, so $\mathbf{a}_1 = [1, 4]$ and $\mathbf{a}_2 = [5, 2]$.

A unit vector maximizing $||A\mathbf{v}||$ is $\mathbf{v}_1 \approx \begin{bmatrix} .78 \\ .63 \end{bmatrix}$.



$\sigma_1 = ||A\mathbf{v}_1||$, which is about 6.1

Theorem

def trolley_line_location(A):
$\mathbf{v}_1 = \arg\max\{||A\mathbf{v}|| \ : \ ||\mathbf{v}|| = 1\}$
$\sigma_1 = ||A\mathbf{v}_1||$
return $\mathbf{v}_1$

Definition: σ_1 is *first singular value* of A.
$\mathbf{v}_1$ is *first right singular vector*.

Theorem: Let A be an $m \times n$ matrix over $\mathbb{R}$ with rows $\mathbf{a}_1, \ldots, \mathbf{a}_m$. Let $\mathbf{v}_1$ be the first right singular vector of A. Then Span $\{\mathbf{v}_1\}$ is the one-dimensional vector space $\mathcal{V}$ that minimizes

$$(\text{distance from } \mathbf{a}_1 \text{ to } \mathcal{V})^2 + \cdots + (\text{distance from } \mathbf{a}_m \text{ to } \mathcal{V})^2$$

How close is the closest vector space to the rows of A?

Lemma: The minimum sum of squared distances is $||A||_F^2 - \sigma_1^2$.
Proof: The distance is $\sum_i ||\mathbf{a}_i||^2 - \sum_i ||\mathbf{a}_i^{||\mathbf{v}}||^2$.
The first sum is $||A||_F^2$.
The second sum is square of $||A\mathbf{v}_1||$, i.e. square of σ_1. QED

Example, continued

Let $A = \begin{bmatrix} 1 & 4 \\ 5 & 2 \end{bmatrix} \Rightarrow \mathbf{a}_1 = [1,4], \mathbf{a}_2 = [5,2]$. **Solution:** $\mathbf{v}_1 \approx \begin{bmatrix} .78 \\ .63 \end{bmatrix}$. Sum of squared distances?

Projection of $\mathbf{a}_1$ orthogonal to $\mathbf{v}_1$:

$$\begin{aligned} \mathbf{a}_1 - \langle \mathbf{a}_1, \mathbf{v}_1 \rangle \, \mathbf{v}_1 &\approx [1,4] - (1 \cdot .78 + 4 \cdot .63)[.78, .63] \\ &\approx [1,4] - 3.3\,[.78, .63] \\ &\approx [-1.6, 1.9] \end{aligned}$$

Norm, about 2.5, is distance from $\mathbf{a}_1$ to Span $\{\mathbf{v}_1\}$.

Projection of $\mathbf{a}_2$ orthogonal to $\mathbf{v}_1$:

$$\begin{aligned} \mathbf{a}_2 - \langle \mathbf{a}_1, \mathbf{v}_1 \rangle \, \mathbf{v}_1 &\approx [5,2] - (5 \cdot .78 + 2 \cdot .63)[.78, .63] \\ &\approx [5,2] - 5.1\,[.78, .63] \\ &\approx [1, -1.2] \end{aligned}$$

Norm, about 1.6, is distance from $\mathbf{a}_2$ to Span $\{\mathbf{v}_1\}$.

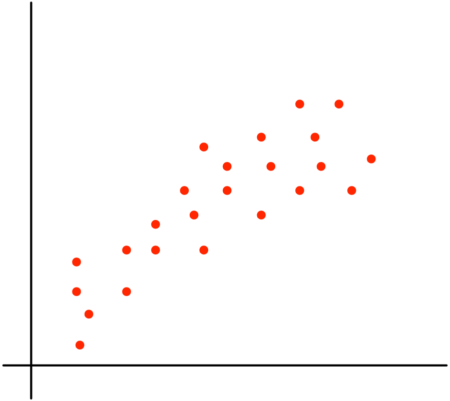
Thus the sum of squared distances is about $2.5^2 + 1.6^2$, which is about 8.7.

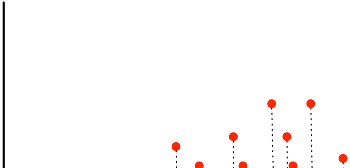
Lemma says sum of squared distances should be

$||A||_F^2 - \sigma_1^2 \approx (1^2 + 4^2 + 5^2 + 2^2) - 6.1^2 \approx 46 - 6.1^2 \approx 8.7$. ✓

Visualization of data in one dimension

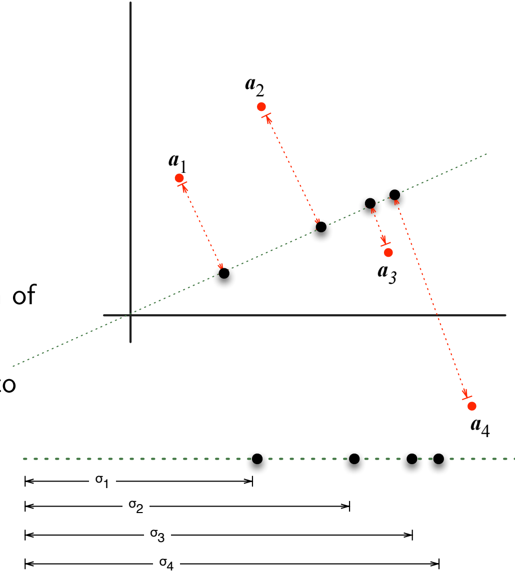
Projections of high-dimensional data points $\mathbf{a}_1, \ldots, \mathbf{a}_m$ onto line: **visualization technique.**





Visualization of data in one dimension

Projections of high-dimensional data points $\mathbf{a}_1, \ldots, \mathbf{a}_m$ onto line: **visualization technique.**



Each datapoint $\mathbf{a}_i$ is represented by a single number:
$\sigma_i = \langle \mathbf{a}_i, \mathbf{v}_1 \rangle$

What do we know about these numbers?

$\mathbf{v}_1$ is chosen among norm-1 vectors to *maximize* the sum of squares of these numbers.

That is, we are choosing a line through the origin so as to maximally *spread out* those numbers.

Application to voting data

Let $\mathbf{a}_1, \ldots, \mathbf{a}_{100}$ be the voting records for US Senators.
Same as you used in politics lab.
These are 46-vectors with ± 1 entries.
Find the unit-norm vector $\mathbf{v}$ that minimizes least-squares distance from $\mathbf{a}_1, \ldots, \mathbf{a}_{100}$ to Span $\{\mathbf{v}\}$.
Look at projection along $\mathbf{v}$ of each of these vectors.

Not so meaningful:

Snowe	0.106605199	moderate Republican from Maine
Lincoln	0.106694552	moderate Republican from Rhode Island
Collins	0.107039376	moderate Republican from Maine
Crapo	0.107259689	not so moderate Republican from Idaho
Vitter	0.108031374	not so moderate Republican from Louisiana

We'll have to come back to this data.

Best rank-one approximation to a matrix

A rank-one matrix is a matrix whose row space is one-dimensional.
All rows must lie in Span $\{\mathbf{v}\}$ for some vector $\mathbf{v}$.
That is, every row is a scalar multiple of $\mathbf{v}$.

outer product $\left[\begin{array}{c} \\ \mathbf{u} \\ \\ \end{array}\right]\left[\begin{array}{ccc} & \mathbf{v} & \end{array}\right]$

Goal: Given matrix A, find the rank-one matrix $\tilde{A}$ that minimizes $\|A - \tilde{A}\|_F$.

$$\tilde{A} = \left[\begin{array}{c} \text{vector in Span } \{\mathbf{v}\} \text{ closest to } \mathbf{a}_1 \\ \hline \vdots \\ \hline \text{vector in Span } \{\mathbf{v}\} \text{ closest to } \mathbf{a}_m \end{array}\right]$$

How close is $\tilde{A}$ to A?

$$\begin{aligned} ||A - \tilde{A}||_F^2 &= \sum_i ||\text{row } i \text{ of } A - \tilde{A}||^2 \\ &= \sum_i ||\text{distance from } \mathbf{a}_i \text{ to Span } \{\mathbf{v}\}||^2 \end{aligned}$$

To minimize the sum of squares of distances, choose $\mathbf{v}$ to be first right singular vector. Sum of squared distances is $||A||_F^2 - \sigma_1^2$.
$\tilde{A}$= closest rank-one matrix.

An expression for the best rank-one approximation

Using the formula $\mathbf{a}_i^{\|\mathbf{v}_1} = \langle \mathbf{a}_i, \mathbf{v}_1 \rangle \, \mathbf{v}_1$, we obtain

$$\tilde{A} = \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{v}_1 \rangle \, \mathbf{v}_1^T \\ \hline \vdots \\ \hline \langle \mathbf{a}_m, \mathbf{v}_1 \rangle \, \mathbf{v}_1^T \end{bmatrix}$$

Using the linear-combinations interpretation of vector-matrix multiplication, we can write this as an outer product of two vectors:

$$\tilde{A} = \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{v}_1 \rangle \\ \vdots \\ \langle \mathbf{a}_m, \mathbf{v}_1 \rangle \end{bmatrix} \begin{bmatrix} \quad \mathbf{v}_1^T \quad \end{bmatrix}$$

The first vector in the outer product can be written as $A\mathbf{v}_1$. We obtain

$$\tilde{A} = \begin{bmatrix} \\ A\mathbf{v}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \quad \mathbf{v}_1^T \quad \end{bmatrix}$$

Remember $\sigma_1 = \|A\mathbf{v}_1\|$. Define $\mathbf{u}_1$ to be the norm-one vector such that $\sigma_1 \, \mathbf{u}_1 = A\mathbf{v}_1$. Then

$$\tilde{A} = \sigma_1 \begin{bmatrix} \\ \mathbf{u}_1 \\ \\ \end{bmatrix} \begin{bmatrix} \quad \mathbf{v}_1^T \quad \end{bmatrix}$$