ENGR-E 511; ENGR-E 399

"Machine Learning for Signal Processing"

Module 01: Lecture 02:

Linear Algebra

Minje Kim

Department of Intelligent Systems Engineering

Email: minje@indiana.edu

Website: http://minjekim.com

Research Group: http://saige.sice.indiana.edu

Meeting Request: http://doodle.com/minje

INDIANA UNIVERSITY

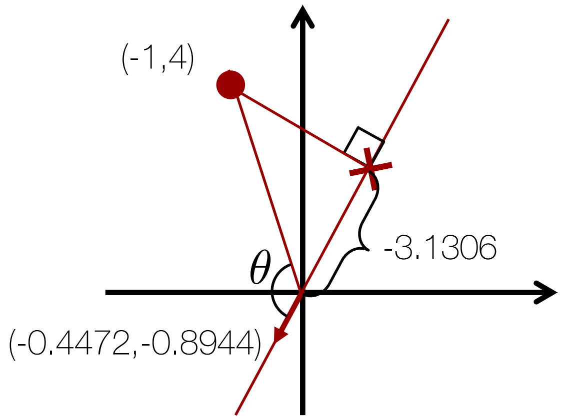
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

Warm-up

- Projection

- We've got a data point $\boldsymbol{x} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$

- and a unit vector $\boldsymbol{v} = \begin{bmatrix} -0.4472 \\ -0.8944 \end{bmatrix}$
  - "Unit vector" means that $\boldsymbol{v}^\top \boldsymbol{v} = 1 \quad (\text{or, } ||\boldsymbol{v}|| = 1)$



- What does this mean? (a.k.a. **inner product**)

$$\boldsymbol{v}^\top \boldsymbol{x} = [-0.4472, -0.8944] \begin{bmatrix} -1 \\ 4 \end{bmatrix} = -3.1306$$

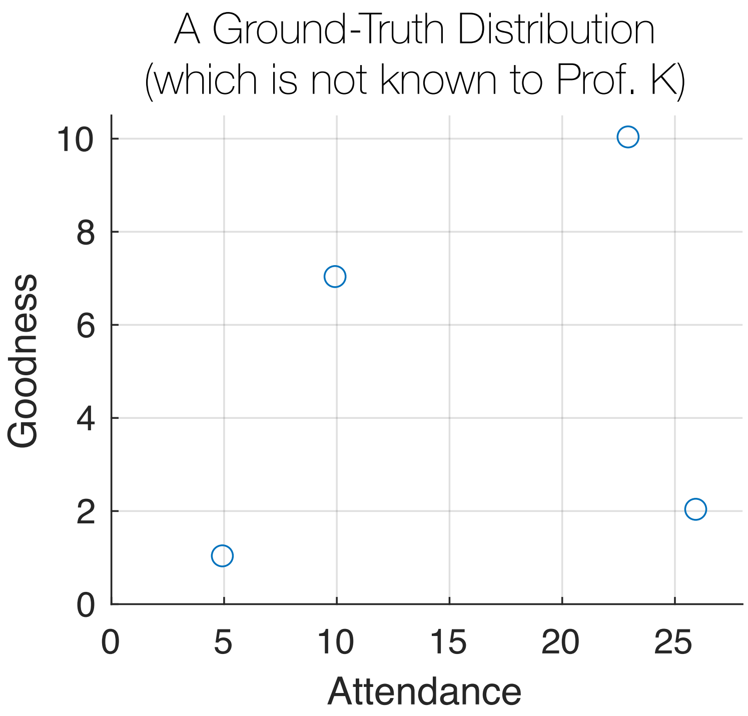
$$= ||\boldsymbol{x}|| \cos(\theta)$$

- **Projection** onto the unit vector produces the new coordinate defined by that unit vector

Inspiring Problem

- Grading

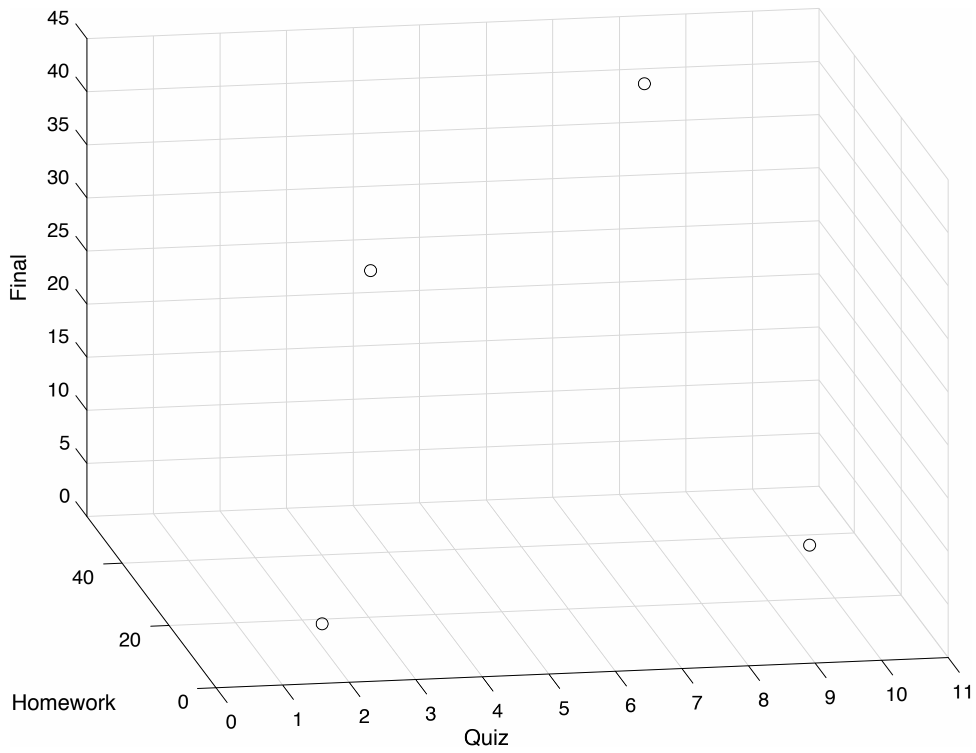
- Prof. K cares about attendance and goodness when it comes to the letter grade
  - But he's too cool to check the attendance
  - And it's difficult to know how good the students are



Inspiring Problem

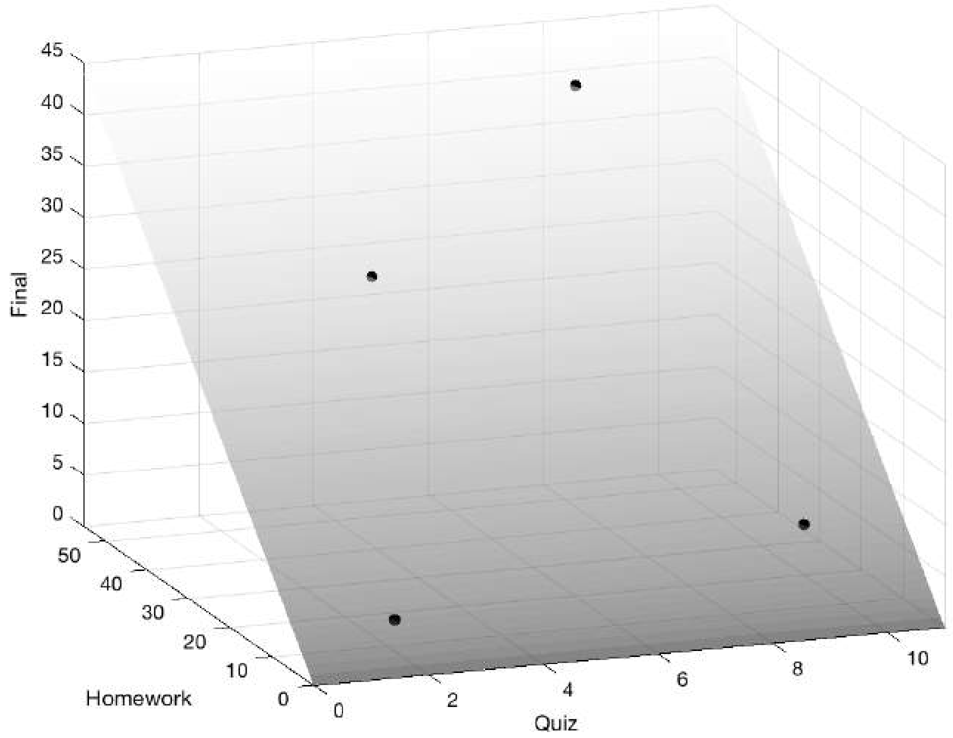
- Grading

- So, he decided to score them indirectly
  - Quiz: 10%, Finals: 40%, Homework: 50% (note that this is different from our class!)



Inspiring Problem
- Three dimensions are redundant

- The data points are with three dimensions, but actually lying on a 2D surface



Inspiring Problem
- Conversion from a space to another

- Suppose two basis vectors that define the subspace of interest

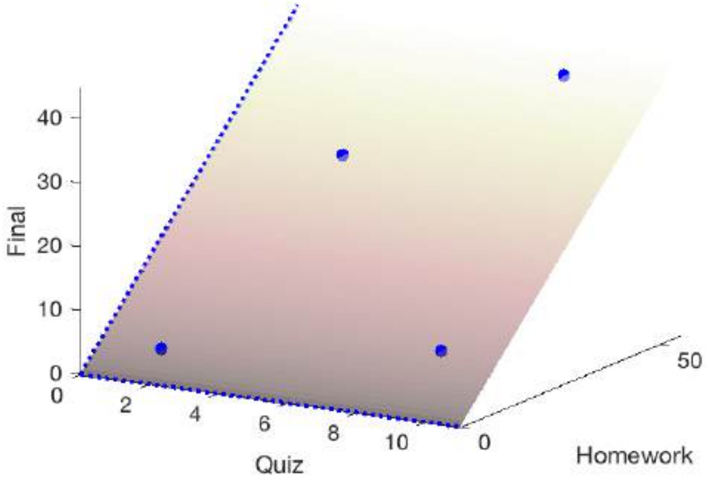
	ST A	ST B	ST C	ST D
Quiz	9.29	3.57	8.21	1.79
Homework	10	35	50	5
Final	8	28	40	4

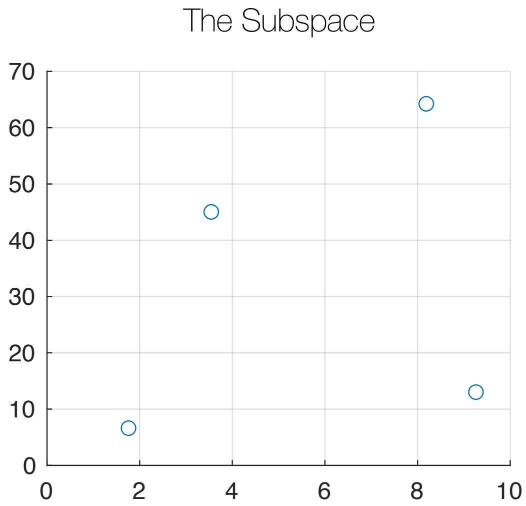
$\approx$

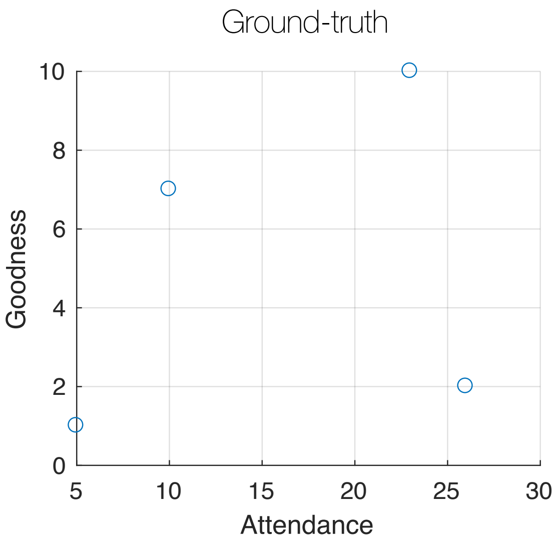
1	0
0	0.78
0	0.62

$\bullet$

9.29	3.57	8.21	1.79
12.81	44.82	64.03	6.40







Inspiring Problem

- Properties of the basis vectors

- Orthonormal!

1	0	0
0	0.78	0.62

●

9.29	3.57	8.21	1.79
10	35	50	5
8	28	40	4

≈

1	0	0
0	0.78	0.62

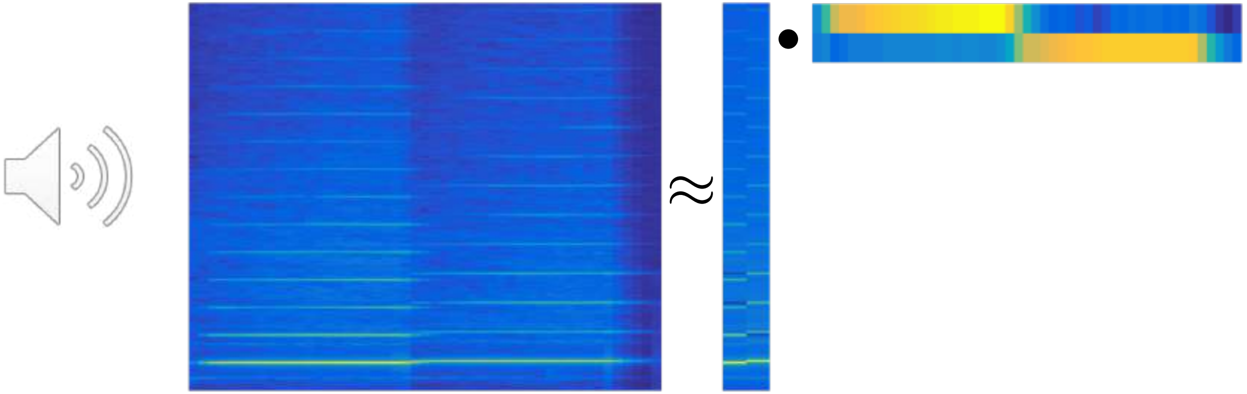
●

1	1	0
0	0	1
0	0.62	

●

9.29	3.57	8.21	1.79
12.81	44.82	64.03	6.40

  - We project the data points onto the basis vectors to get their coordinates in the new space



  - How do we get the orthonormal basis vectors?

Singular Value Decomposition

- Basic definition

- Let's define SVD first:

$$\boldsymbol{A} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_D] \cdot \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_D \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{u}_1^\top \\ \boldsymbol{u}_2^\top \\ \vdots \\ \boldsymbol{u}_D^\top \end{bmatrix}$$

  - Let $s_d = ||\boldsymbol{v}_d^\top \boldsymbol{A}||$ and $\boldsymbol{u}_d = \boldsymbol{A}^\top \boldsymbol{v}_d / s_d$
  - IOW, if you have whatever set of (left) singular vectors, $[\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_D]$ you can define the singular values and the other (right) singular vectors

- Then,

$$\boldsymbol{A} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_D] \cdot \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_D \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{u}_1^\top \\ \boldsymbol{u}_2^\top \\ \vdots \\ \boldsymbol{u}_D^\top \end{bmatrix} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_D] \cdot \begin{bmatrix} \boldsymbol{v}_1^\top \\ \boldsymbol{v}_2^\top \\ \vdots \\ \boldsymbol{v}_D^\top \end{bmatrix} \cdot \boldsymbol{A}$$

This should be an identity matrix

$$\begin{bmatrix} \boldsymbol{v}_1^\top \\ \boldsymbol{v}_2^\top \\ \vdots \\ \boldsymbol{v}_D^\top \end{bmatrix} \cdot \boldsymbol{A} = \boxed{\begin{bmatrix} \boldsymbol{v}_1^\top \\ \boldsymbol{v}_2^\top \\ \vdots \\ \boldsymbol{v}_D^\top \end{bmatrix} \cdot [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_D]} \cdot \begin{bmatrix} \boldsymbol{v}_1^\top \\ \boldsymbol{v}_2^\top \\ \vdots \\ \boldsymbol{v}_D^\top \end{bmatrix} \cdot \boldsymbol{A}$$

  - Therefore, SVD should be with orthonormal basis vectors $\boldsymbol{v}_d$

- It will be nice if it works for $\boldsymbol{A}^\top$, too

$$\boldsymbol{A}^\top = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_D] \cdot \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_D \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{v}_1^\top \\ \boldsymbol{v}_2^\top \\ \vdots \\ \boldsymbol{v}_D^\top \end{bmatrix}$$

  - Therefore, $\boldsymbol{u}_d$ needs to be orthonormal, too

Singular Value Decomposition

- Don’t like math?

- You better… But, let me help you

$$\begin{bmatrix} 9.29 & 3.57 & 8.21 & 1.79 \\ 10 & 35 & 50 & 5 \\ 8 & 28 & 40 & 4 \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ 0 & 0.78 \\ 0 & 0.62 \end{bmatrix} \bullet \begin{bmatrix} 9.29 & 3.57 & 8.21 & 1.79 \\ 12.81 & 44.82 & 64.03 & 6.40 \end{bmatrix}$$

Normalization

$$\begin{bmatrix} 9.29 & 3.57 & 8.21 & 1.79 \\ 10 & 35 & 50 & 5 \\ 8 & 28 & 40 & 4 \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ 0 & 0.78 \\ 0 & 0.62 \end{bmatrix} \bullet \begin{bmatrix} 13.02 & 0 \\ 0 & 79.46 \end{bmatrix} \bullet \begin{bmatrix} 0.71 & 0.27 & 0.63 & 0.14 \\ 0.16 & 0.56 & 0.81 & 0.08 \end{bmatrix}$$

- Is this SVD?
  - No.. Because the right singular vectors are not orthonormal

$$\begin{bmatrix} 0.71 & 0.27 & 0.63 & 0.14 \\ 0.16 & 0.56 & 0.81 & 0.08 \end{bmatrix} \bullet \begin{bmatrix} 0.71 & 0.16 \\ 0.27 & 0.56 \\ 0.63 & 0.81 \\ 0.14 & 0.08 \end{bmatrix} = \begin{bmatrix} 1 & 0.79 \\ 0.79 & 1 \end{bmatrix}$$

- In SVD, we find orthonormal $\boldsymbol{v}_d$ that makes $\boldsymbol{u}_d = \boldsymbol{A}^\top \boldsymbol{v}_d / s_d$ orthonormal, too

Eigendecomposition

- From SVD to eigendecomposition

- Let's start by an eigendecomposition problem
  - Two distinct eigenvalues: $\boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$ $\qquad \boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j$
  - First, $\boldsymbol{A}\boldsymbol{A}^\top$ is positive definite. Why?

$$\boldsymbol{v}_i^\top \boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{v}_i = \lambda_i > 0$$

- Then, since they are distinct $\lambda_i \neq \lambda_j$

$$\boldsymbol{v}_j^\top \boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{v}_i = \lambda_i \boldsymbol{v}_j^\top \boldsymbol{v}_i \qquad \boldsymbol{v}_i^\top \boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{v}_j = \lambda_j \boldsymbol{v}_i^\top \boldsymbol{v}_j$$

$$\boldsymbol{v}_j^\top \boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{v}_i - \boldsymbol{v}_i^\top \boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{v}_j = 0 = (\lambda_i - \lambda_j)\boldsymbol{v}_i^\top \boldsymbol{v}_j$$ ← True only if basis vectors are orthogonal

- We can find **orthogonal vectors** using eigendecomposition
  - Then what?
  - First, now that we know $\boldsymbol{v}_i^\top \boldsymbol{v}_j = 0$
  - Therefore,

$$\boldsymbol{v}_i^\top \boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{v}_j = \lambda_j \boldsymbol{v}_i^\top \boldsymbol{v}_j = 0$$

$$(s_i \boldsymbol{u}_i^\top)(s_j \boldsymbol{u}_j) = 0$$ ←From definition $\qquad \boldsymbol{u}_i = \boldsymbol{A}^\top \boldsymbol{v}_i / s_i$

$$\therefore \boldsymbol{u}_i^\top \boldsymbol{u}_j = 0 \quad (\text{for } i \neq j)$$

Eigendecomposition

- Power iteration

- For a square matrix $\boldsymbol{X}$
  - We call the non-zero vector $\mathbf{v}$ an **eigenvector**
  - λ is an **eigenvalue**
  - If they meet the following equation: $\boldsymbol{X}\mathbf{v} = \lambda\mathbf{v}$
  - Or: $\boldsymbol{X} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1} \quad (\boldsymbol{\Lambda} = \mathrm{diag}([\lambda_1, \lambda_2, \cdots, \lambda_D]), \quad \lambda_1 > \lambda_2 > \cdots > \lambda_D)$
- **Power Iteration**

$$\mathbf{v}_1 \approx \lim_{k \to \infty} \boldsymbol{y}^{(k)}$$

$$\boldsymbol{y}^{(k)} = \frac{\boldsymbol{X}\boldsymbol{y}^{(k-1)}}{||\boldsymbol{X}\boldsymbol{y}^{(k-1)}||}$$

$$\begin{aligned}
\boldsymbol{X}^k\boldsymbol{y} &= (\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1})^k\boldsymbol{y} = \boldsymbol{V}\boldsymbol{\Lambda}^k\boldsymbol{V}^{-1}\boldsymbol{y} \\
&= \lambda_1^k\boldsymbol{V}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \left(\frac{\lambda_2}{\lambda_1}\right)^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\frac{\lambda_D}{\lambda_1}\right)^k \end{bmatrix}\boldsymbol{V}^{-1}\boldsymbol{y} \\
&\approx \lambda_1^k\boldsymbol{V}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}\boldsymbol{V}^{-1}\boldsymbol{y} \qquad (\text{ for a large } k) \\
&= \lambda_1^k[\mathbf{v}_1, 0, \cdots, 0]\cdot\tilde{\boldsymbol{y}} \qquad (\tilde{\boldsymbol{y}} = \boldsymbol{V}^{-1}\boldsymbol{y}) \\
&= \lambda_1^k\tilde{y}_1\mathbf{v}_1
\end{aligned}$$

Eigendecomposition

- Remaining problems

- What about the other eigenvectors?

$$\boldsymbol{A} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_D] \cdot \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_D \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{u}_1^\top \\ \boldsymbol{u}_2^\top \\ \vdots \\ \boldsymbol{u}_D^\top \end{bmatrix}$$

$$\boldsymbol{A} = \boldsymbol{v}_1 s_1 \boldsymbol{u}_1^\top + [\boldsymbol{v}_2, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_D] \cdot \begin{bmatrix} s_2 & 0 & \cdots & 0 \\ 0 & s_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_D \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{u}_2^\top \\ \boldsymbol{u}_3^\top \\ \vdots \\ \boldsymbol{u}_D^\top \end{bmatrix}$$

  - We can find the singular vectors one-by-one in the order of eigenvalues

Eigendecomposition

- Eigenfaces

Less eigenfaces

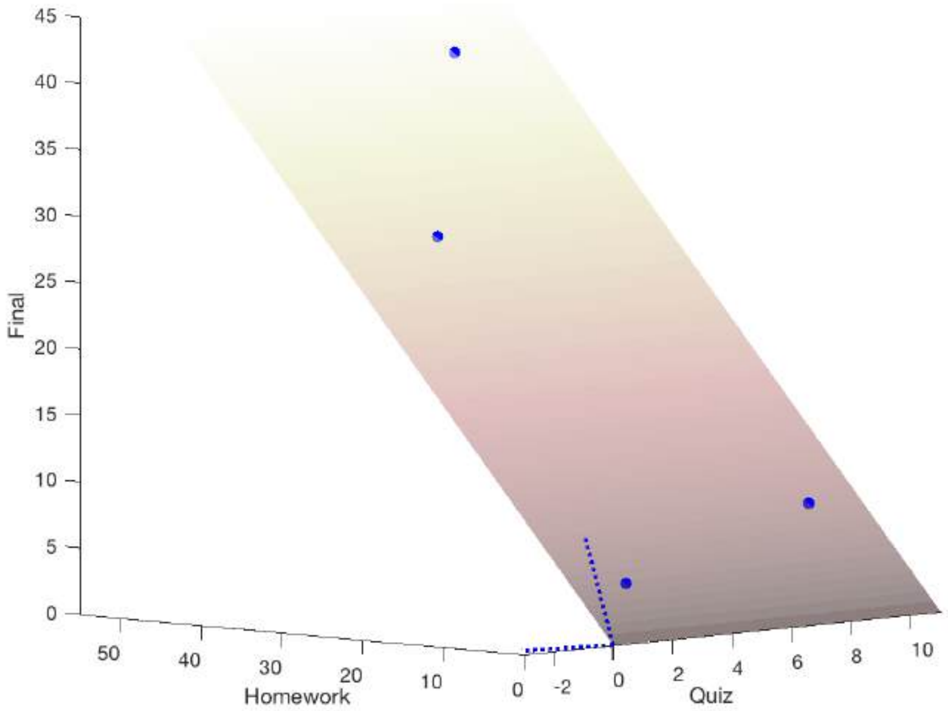
More eigenfaces

http://docs.opencv.org/2.4/_images/eigenface_reconstruction_opencv.png

Eigendecomposition

- Singular vectors from power iteration

- MATLAB figure



Open questions

- Do basis vectors have to be orthogonal?
- What is the meaning behind the order of eigenvalues?

Reading Material

- Textbook Appendix B
- A great demo about SVD: http://websites.uwlax.edu/twill/svd/

Thank You!