Principal Component Analysis Demystified

Caroline Walker

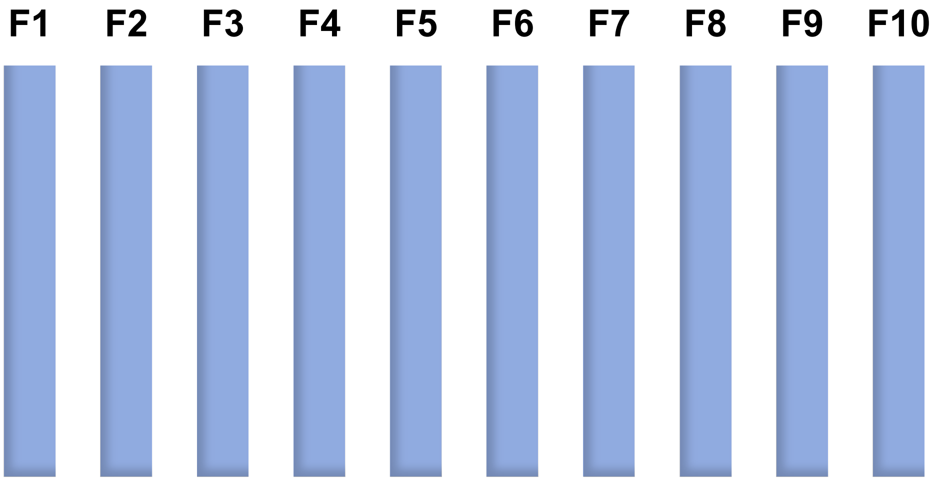
Warren Rogers LLC

Principal Component Analysis

Goal: Dimensionality Reduction

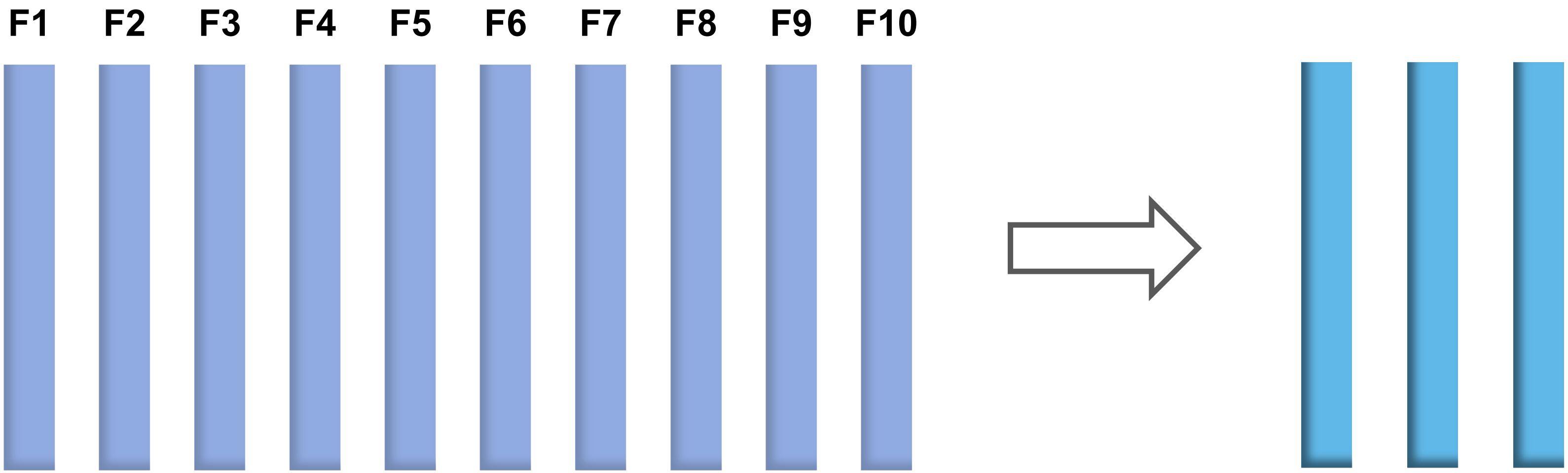
Principal Component Analysis

Goal: **Dimensionality Reduction**



Principal Component Analysis

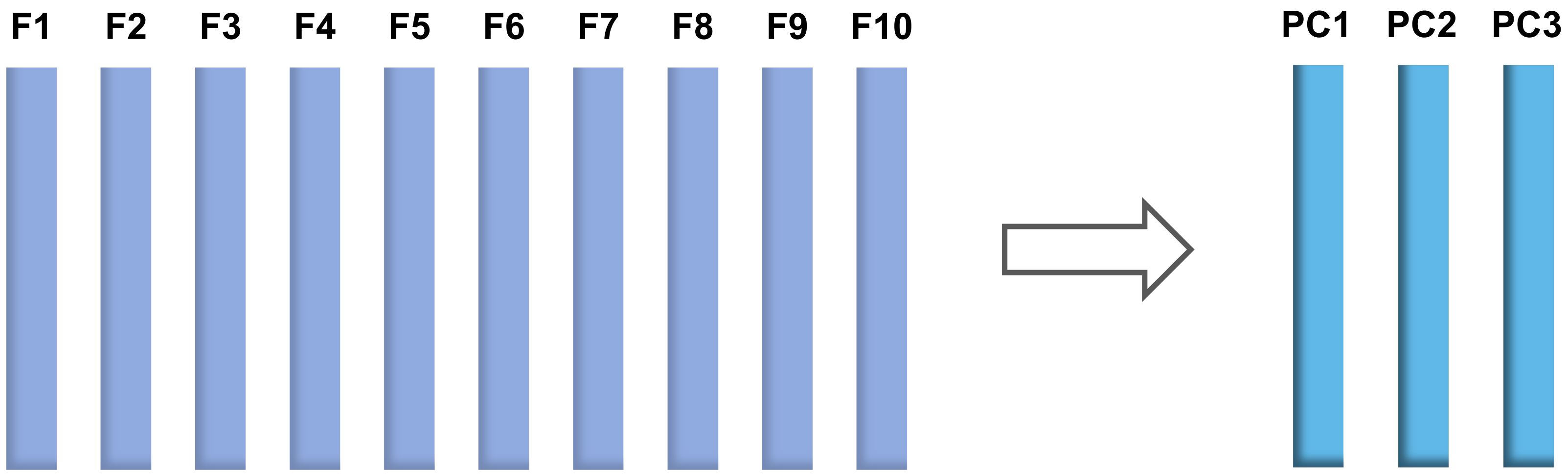
Goal: Dimensionality Reduction



#SASGF

Principal Component Analysis

Goal: Dimensionality Reduction



Principal Component Analysis

The new features…

Principal Component Analysis

The new features…

- Are linear transformations of the original features

Principal Component Analysis

The new features…

- Are linear transformations of the original features
- Are linearly uncorrelated with each other

Principal Component Analysis

The new features…

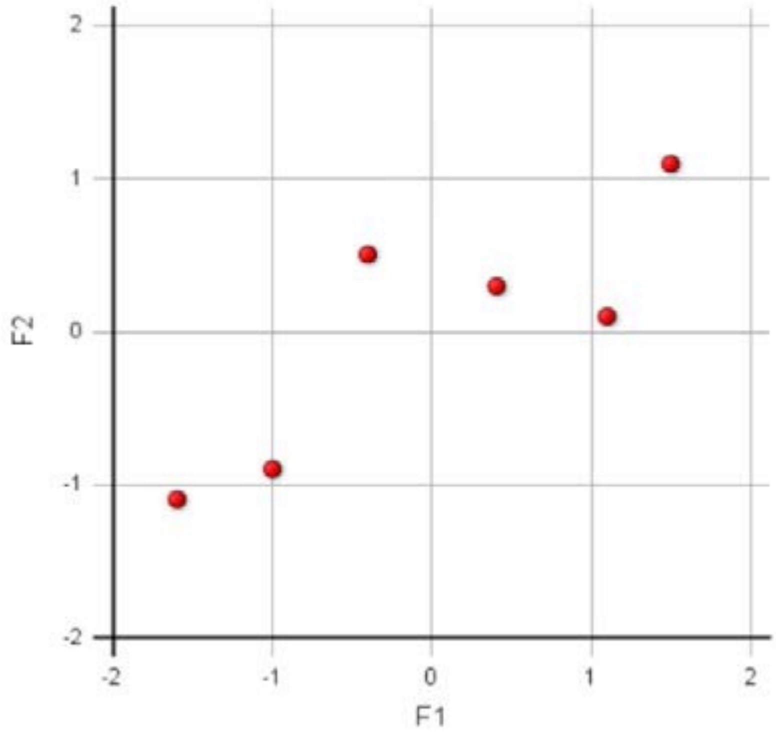
- Are linear transformations of the original features
- Are linearly uncorrelated with each other
- Retain maximum variance from the original feature set

Principal Component Analysis

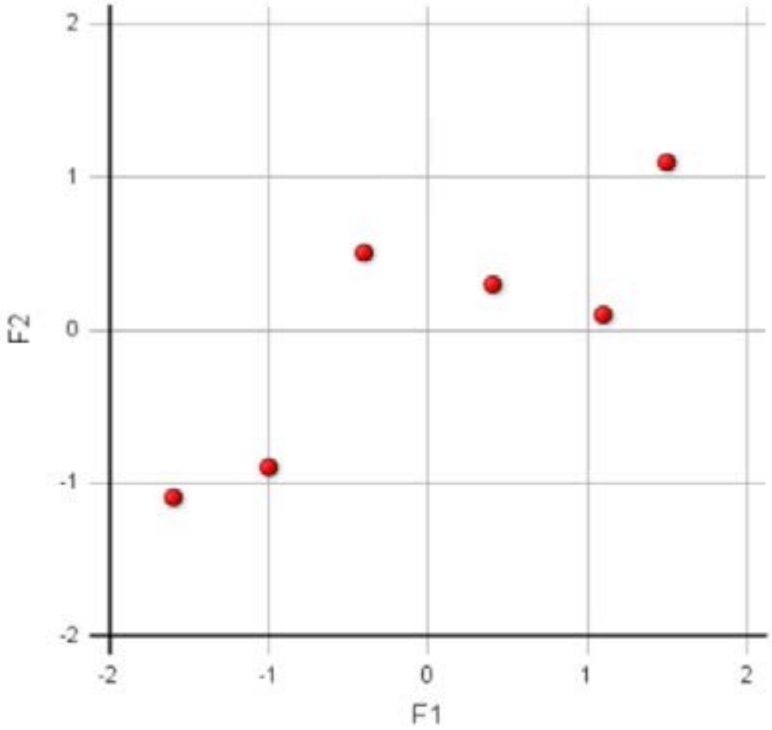
The new features…

- Are linear transformations of the original features
- Are linearly uncorrelated with each other
- Retain maximum variance from the original feature set

An Example in Two Dimensions

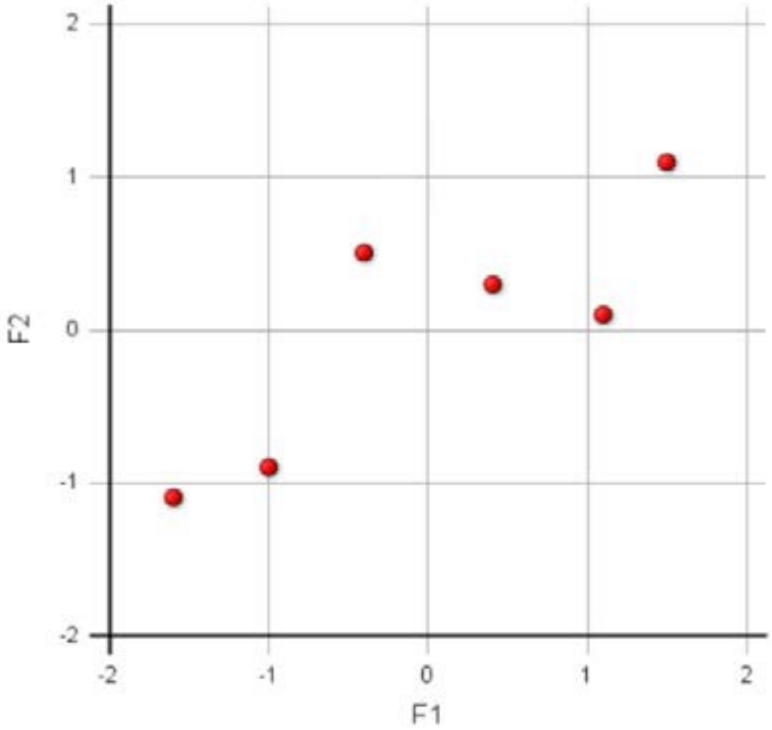


An Example in Two Dimensions



Feature	Variance
F1	1.468
F2	0.716

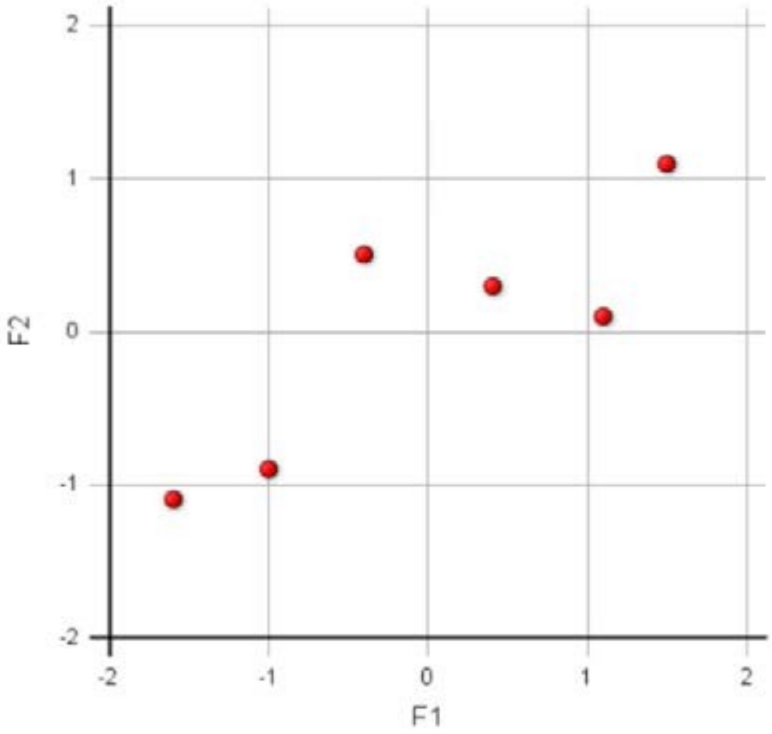
An Example in Two Dimensions



Feature	Variance
F1	1.468
F2	0.716

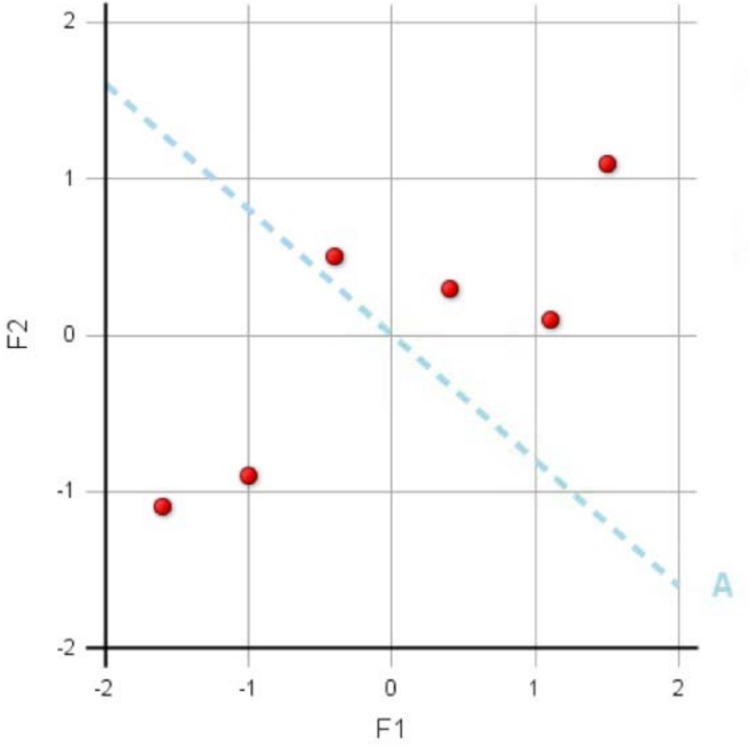
2.184

An Example in Two Dimensions



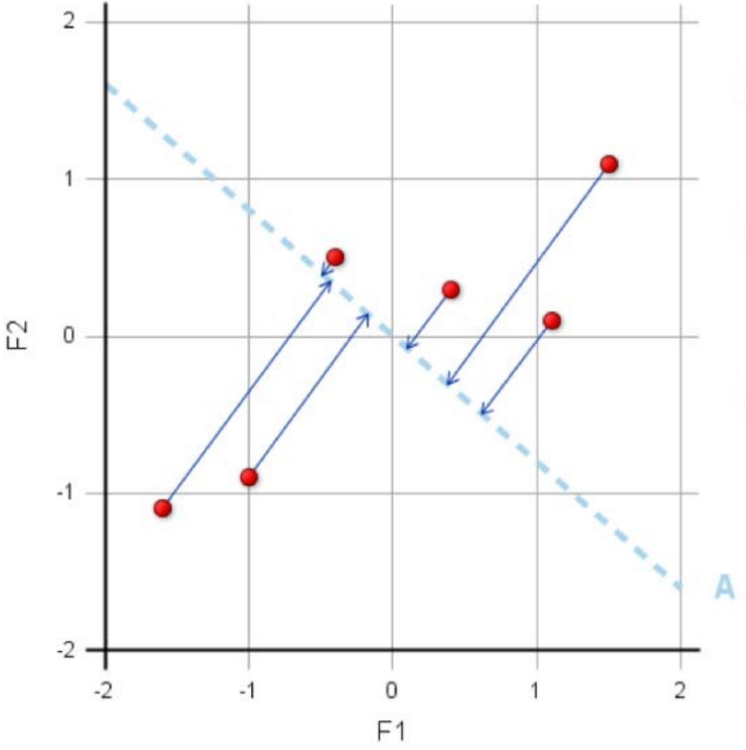
Evaluating a Possible New Feature

Feature A



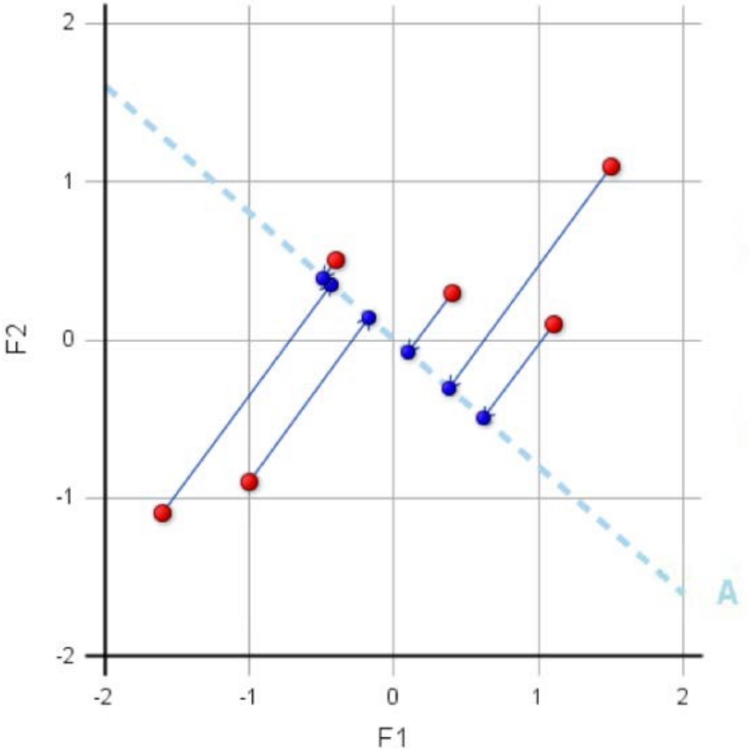
Evaluating a Possible New Feature

Feature A



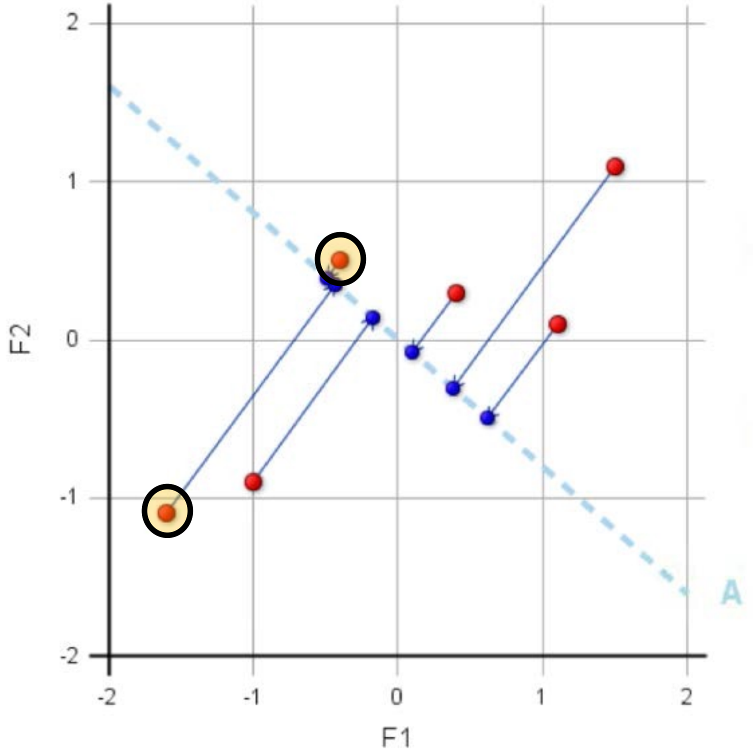
Evaluating a Possible New Feature

Feature A



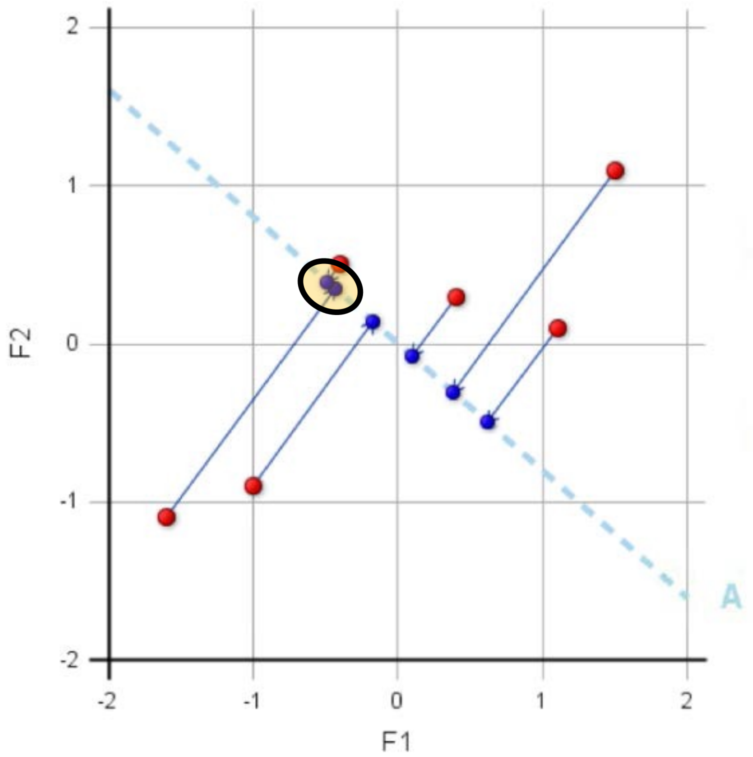
Evaluating a Possible New Feature

Feature A



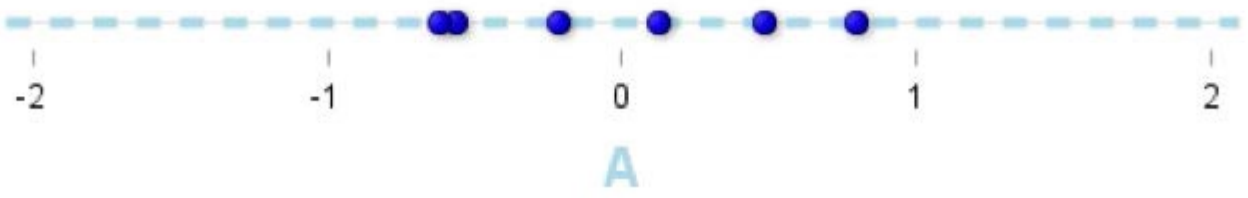
Evaluating a Possible New Feature

Feature A



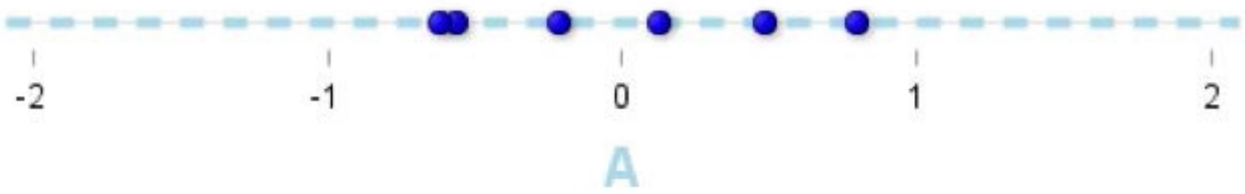
Evaluating a Possible New Feature

Feature A



Evaluating a Possible New Feature

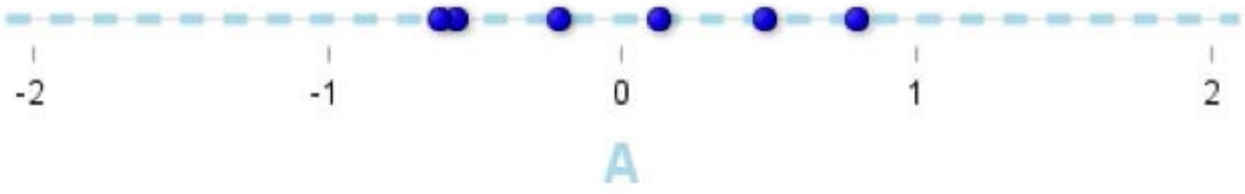
Feature A



Feature	Variance
A	0.328

Evaluating a Possible New Feature

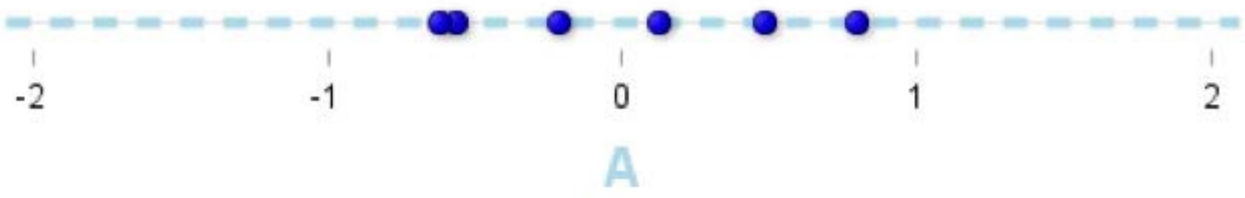
Feature A



Feature	Variance
A	**0.328**
F1	1.468
F2	0.716

Evaluating a Possible New Feature

Feature A

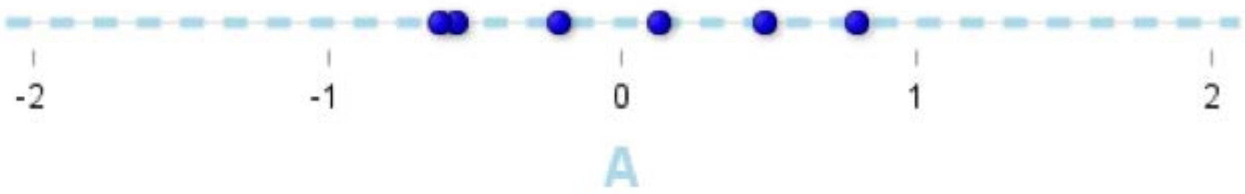


Feature	Variance
A	**0.328**
F1	1.468
F2	0.716

2.184

Evaluating a Possible New Feature

Feature A



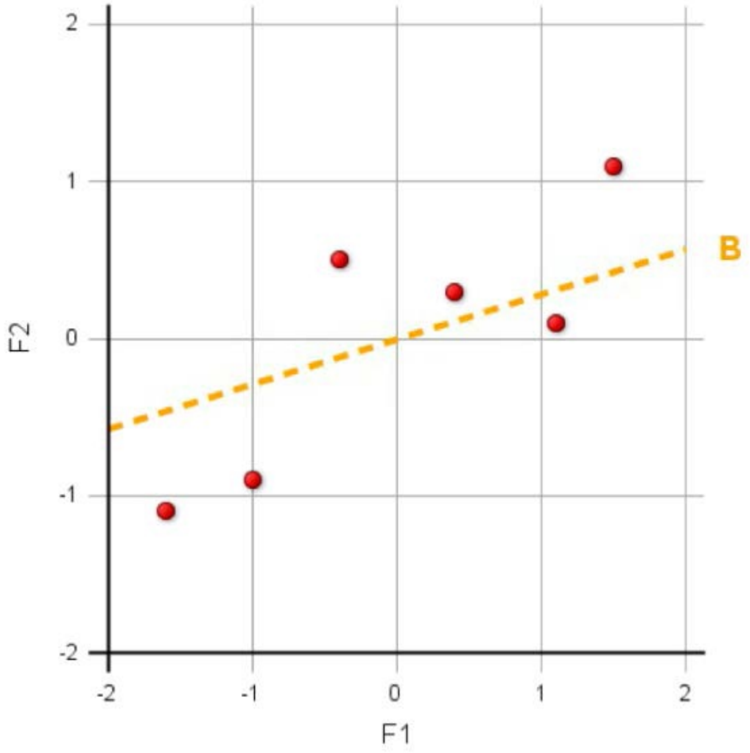
Feature	Variance
A	0.328
F1	1.468
F2	0.716

15%

2.184

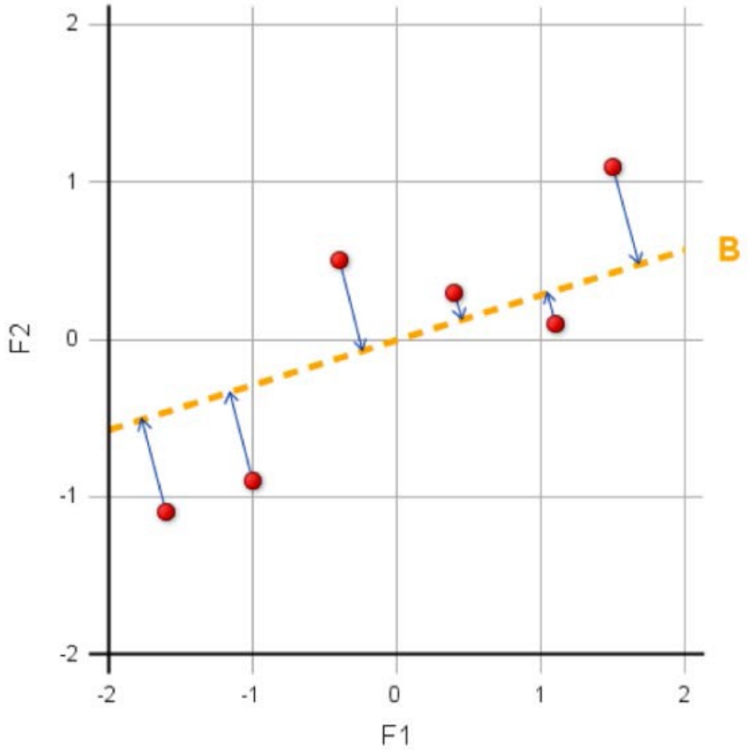
Evaluating a Possible New Feature

Feature B



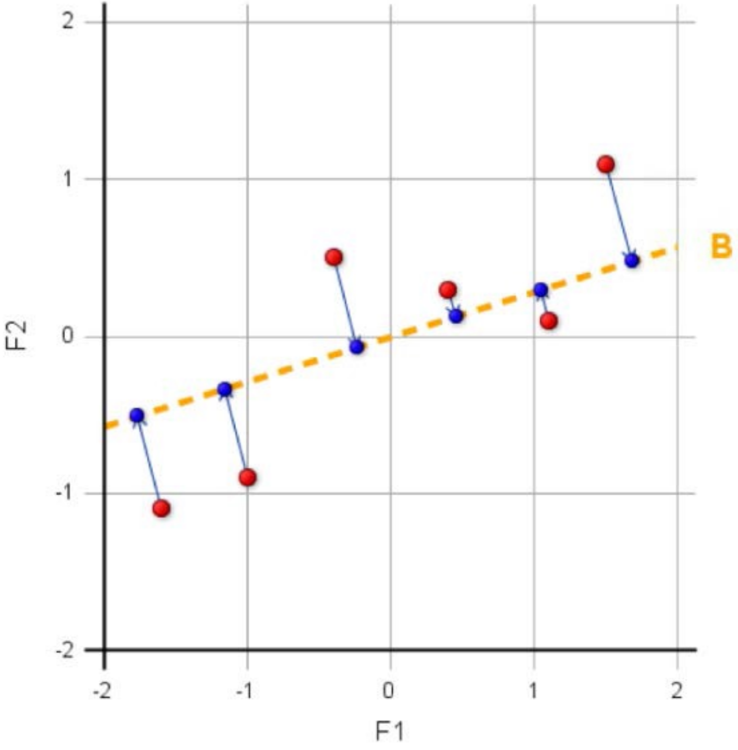
Evaluating a Possible New Feature

Feature B



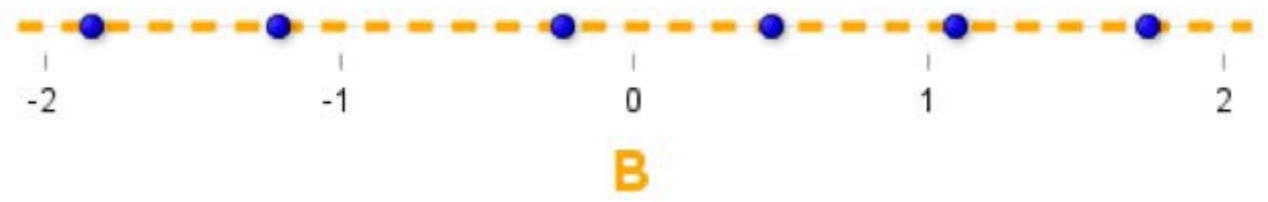
Evaluating a Possible New Feature

Feature B



Evaluating a Possible New Feature

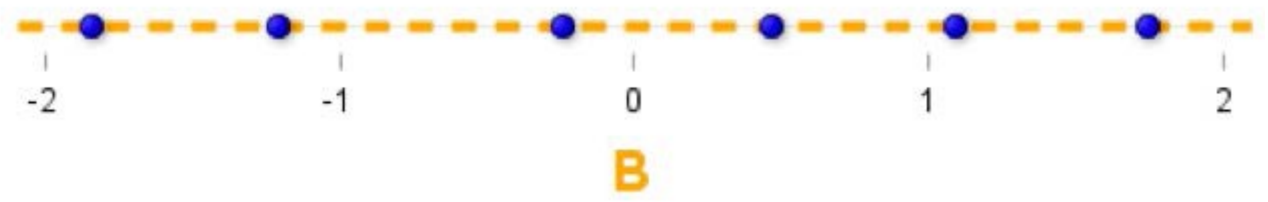
Feature B



Feature	Variance
B	1.869

Evaluating a Possible New Feature

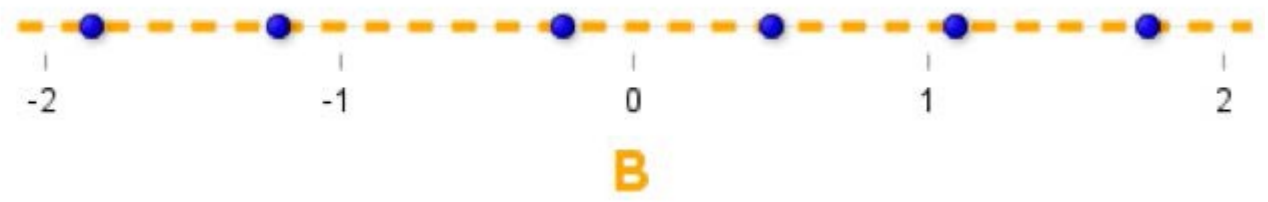
Feature B



Feature	Variance
B	1.869
F1	1.468
F2	0.716

Evaluating a Possible New Feature

Feature B

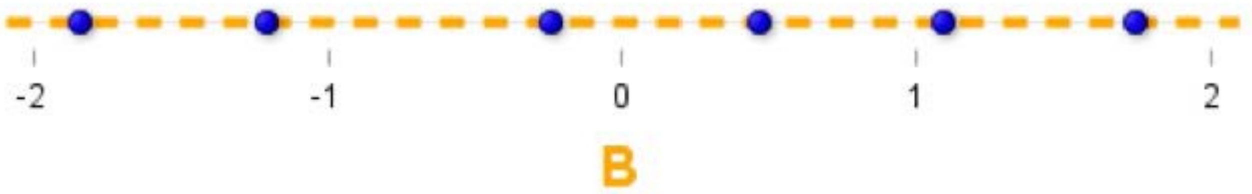


Feature	Variance
B	1.869
F1	1.468
F2	0.716

2.184

Evaluating a Possible New Feature

Feature B



Feature	Variance
B	1.869
F1	1.468
F2	0.716

86%

$\underbrace{1.468 + 0.716}_{}$ 2.184

Finding the Optimal Solution

Steps of PCA

Finding the Optimal Solution

Steps of PCA

1. Calculate the covariance* matrix of the original feature set.

Finding the Optimal Solution

Steps of PCA

1. Calculate the covariance* matrix of the original feature set.

2. Find the eigenvectors and eigenvalues of this covariance matrix.

Finding the Optimal Solution

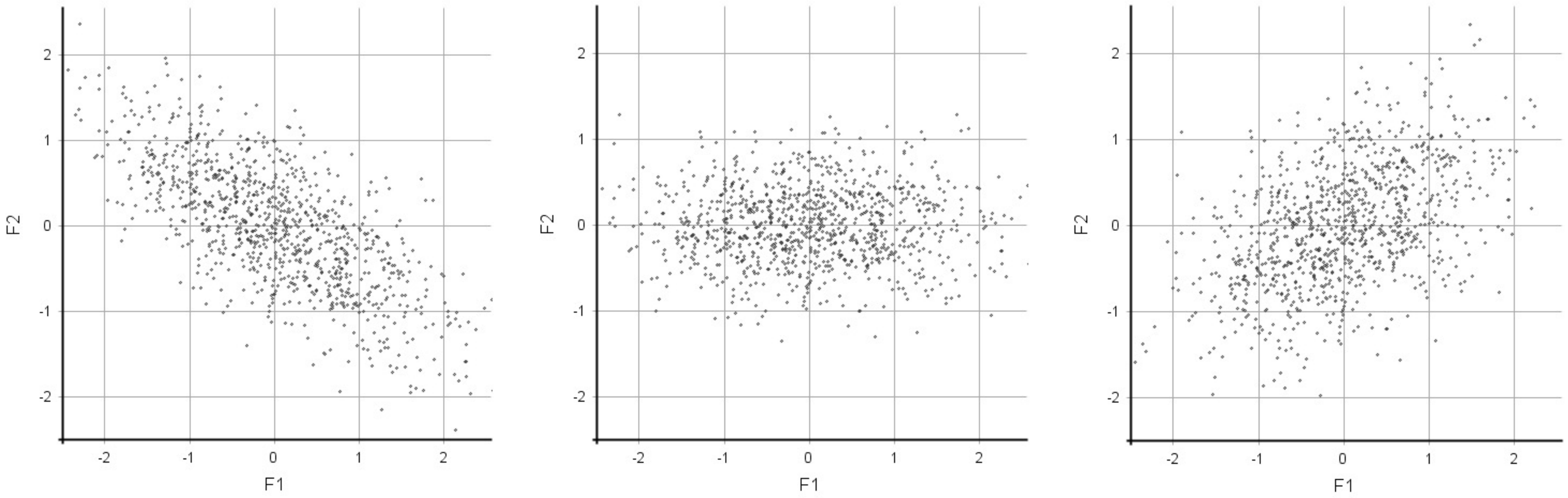
Steps of PCA

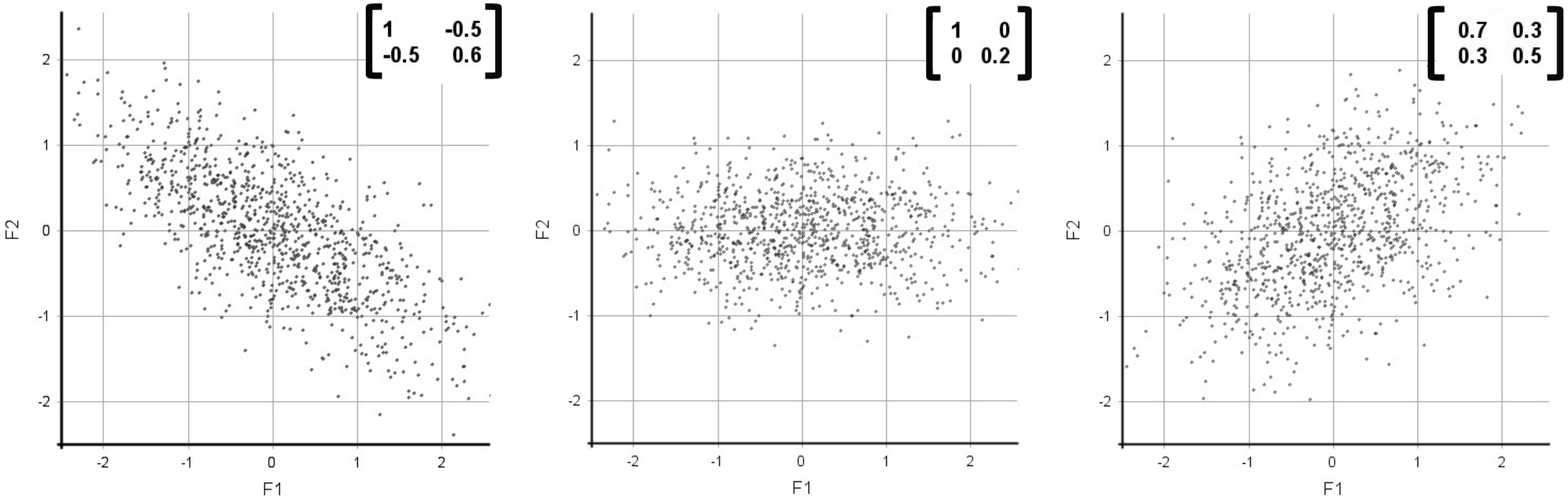
1. Calculate the covariance* matrix of the original feature set.
2. Find the eigenvectors and eigenvalues of this covariance matrix.
3. Order the eigenvectors according to the magnitude of their eigenvalues.

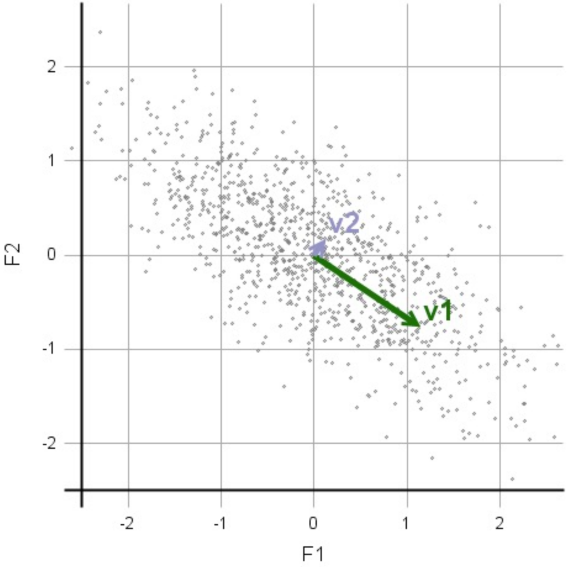
Finding the Optimal Solution

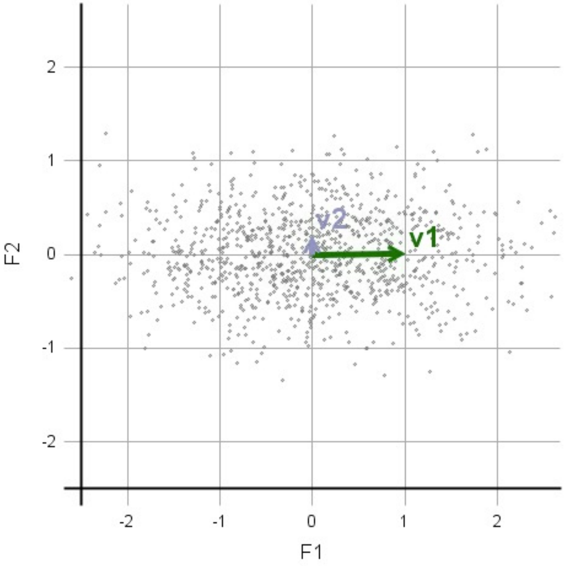
Steps of PCA

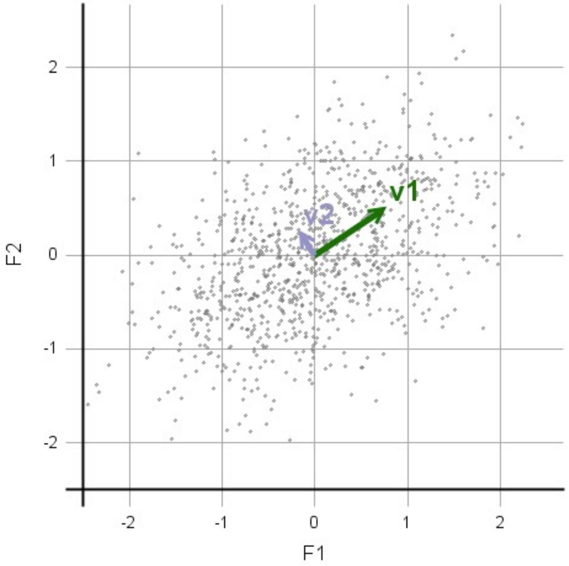
1. Calculate the covariance* matrix of the original feature set.
2. Find the eigenvectors and eigenvalues of this covariance matrix.
3. Order the eigenvectors according to the magnitude of their eigenvalues.
4. **The eigenvector v_1 will show the direction of maximum variance within the data set.**

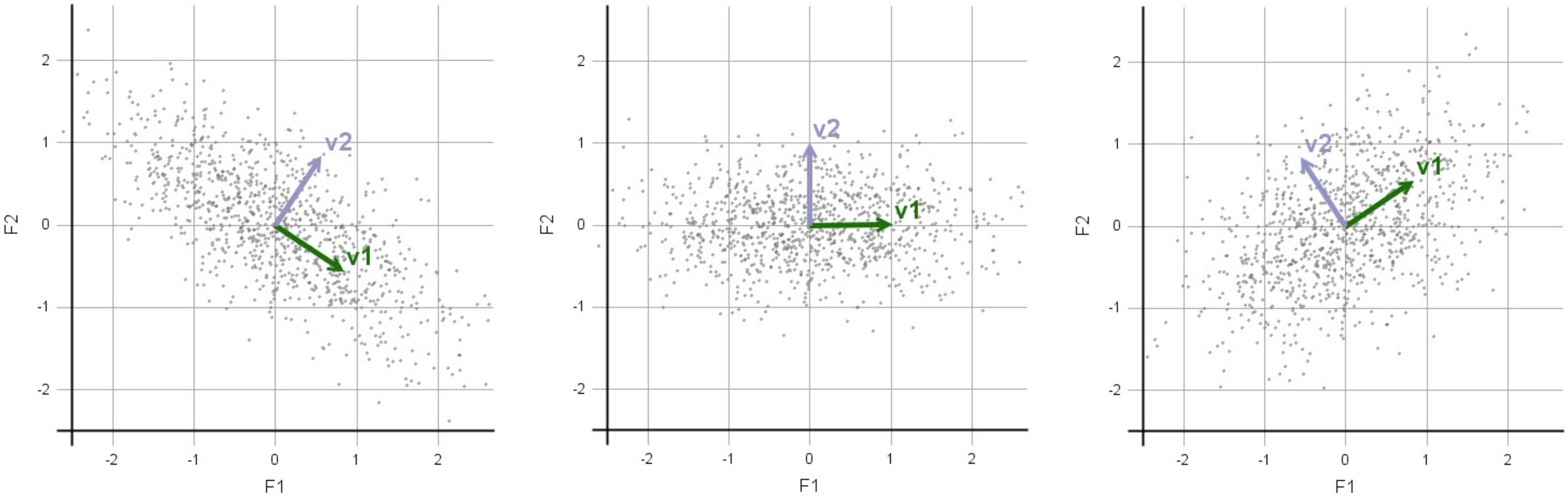

F2
F1
2
1
0
-1
-2
-2
-1
0
1
2


1 -0.5
-0.5 0.6
1 0
0 0.2
0.7 0.3
0.3 0.5
F1
F2
2
1
0
-1
-2


F2
F1
v1
v2
2
1
0
-1
-2
-2
-1
0
1
2

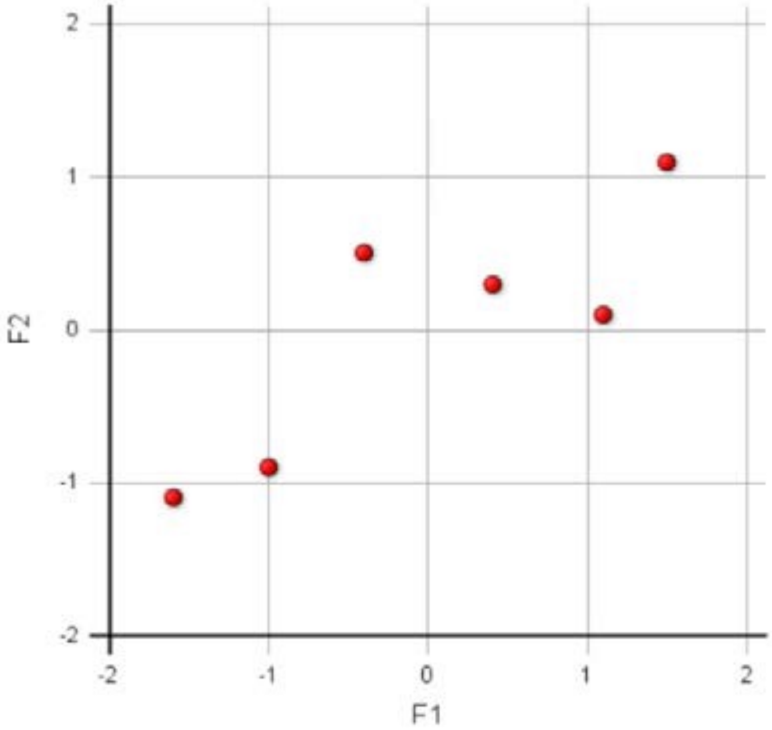

F2
F1
v1
v2
2
1
0
-1
-2
-2
-1
0
1
2


F2
F1
v1
v2
2
1
0
-1
-2
-2
-1
0
1
2



v2
v1
F1
F2
2
1
0
-1
-2

Principal Component Analysis



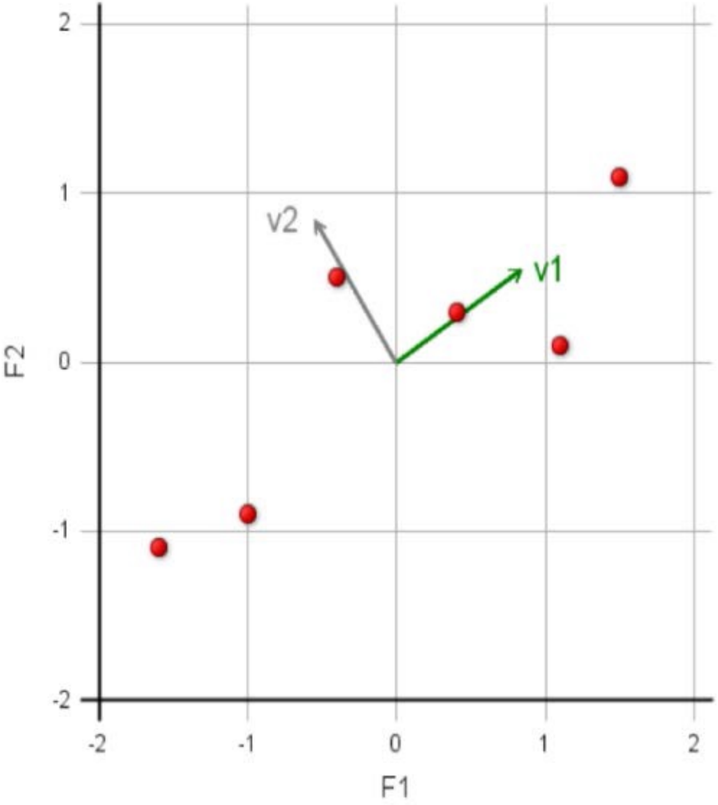
$$\begin{bmatrix} 1.468 & 0.868 \\ 0.868 & 0.716 \end{bmatrix}$$

Principal Component Analysis

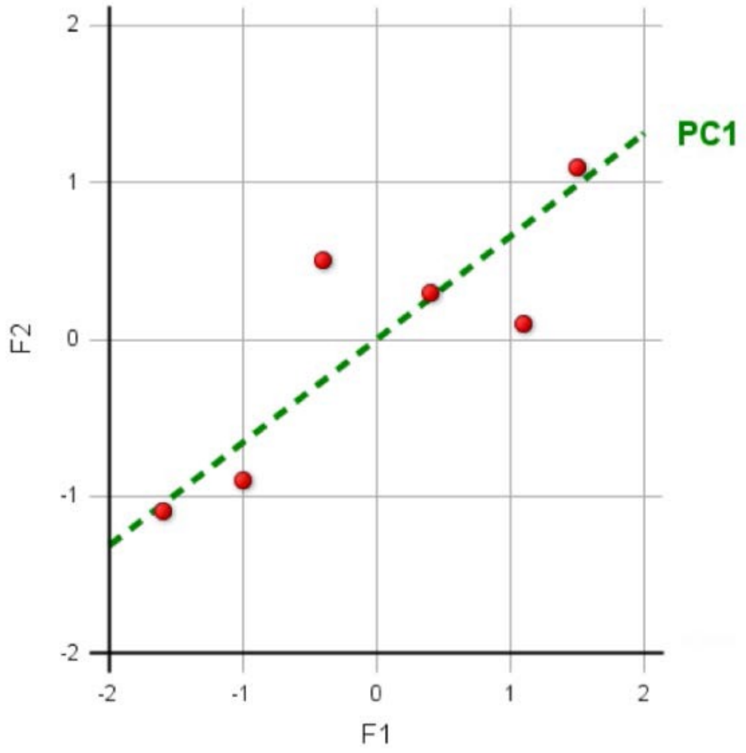
$$\begin{bmatrix} 0.836 \\ 0.549 \end{bmatrix} \qquad \begin{bmatrix} -0.549 \\ 0.836 \end{bmatrix}$$

$$\lambda_1 = 2.038 \qquad \lambda_2 = 0.146$$

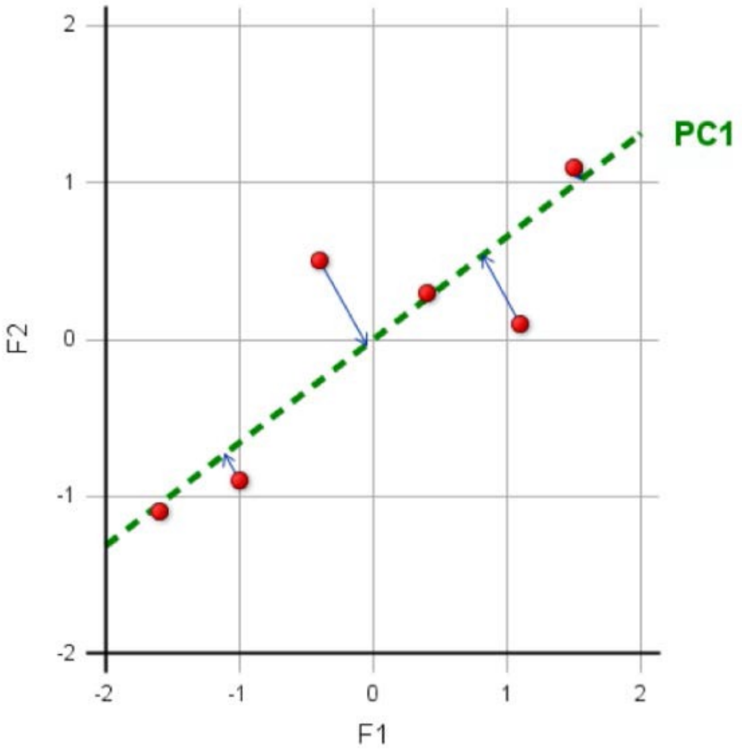
Principal Component Analysis



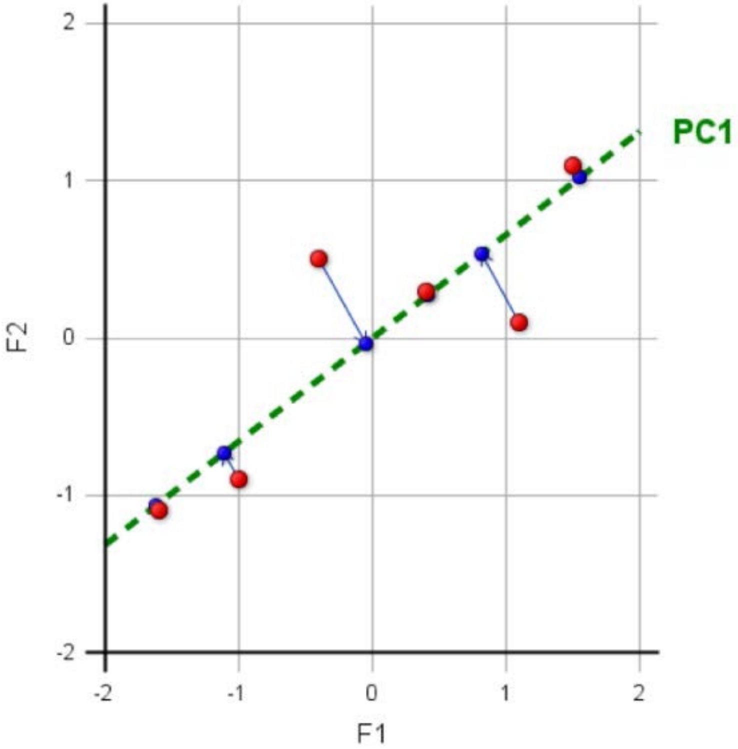
Principal Component Analysis



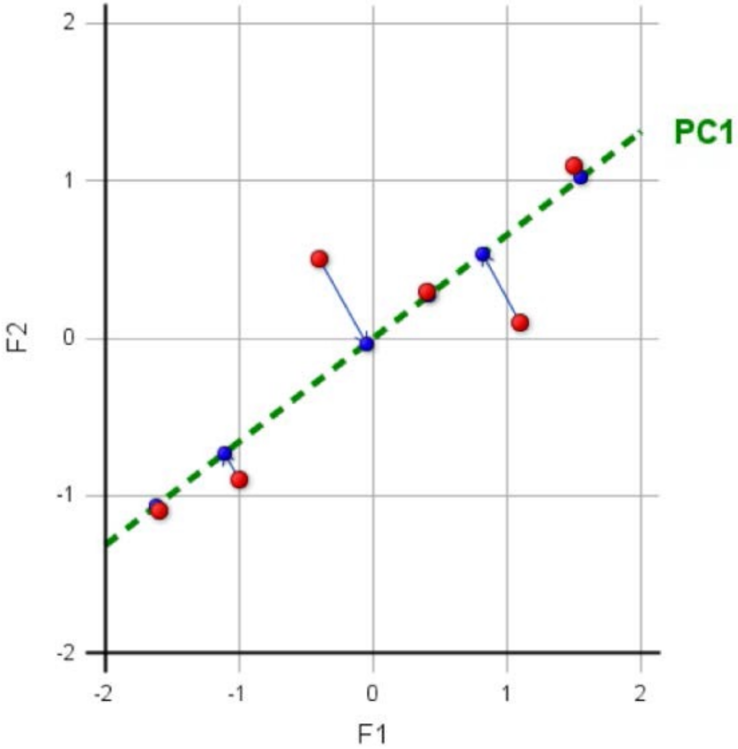
Principal Component Analysis



Principal Component Analysis

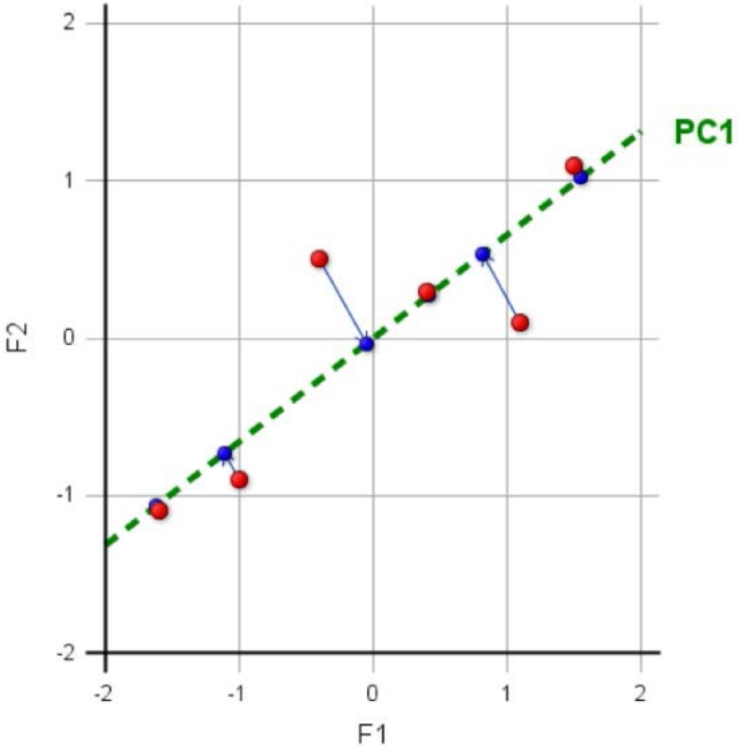


Principal Component Analysis



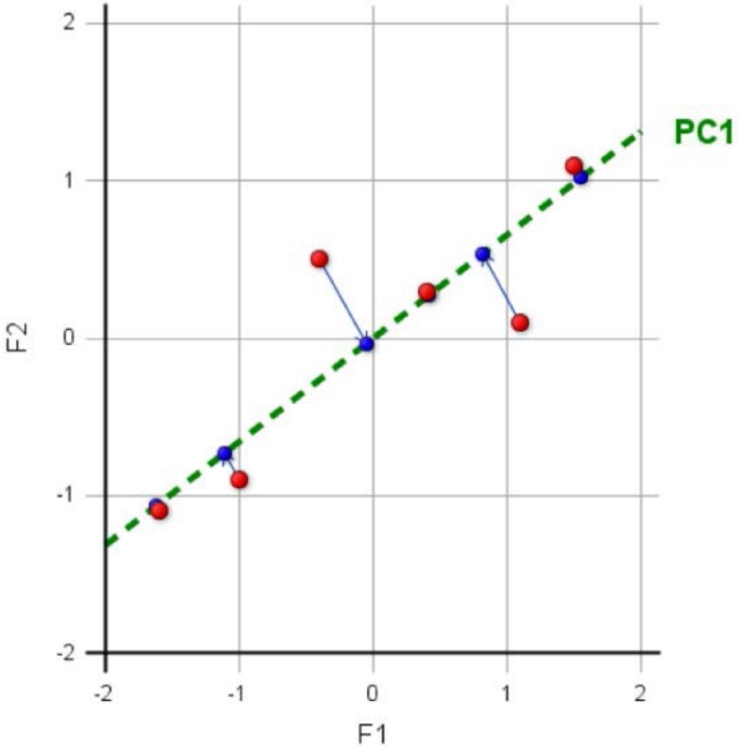
Feature	Variance
PC1	2.038

Principal Component Analysis



Feature	Variance
PC1	2.038
F1	1.468
F2	0.716

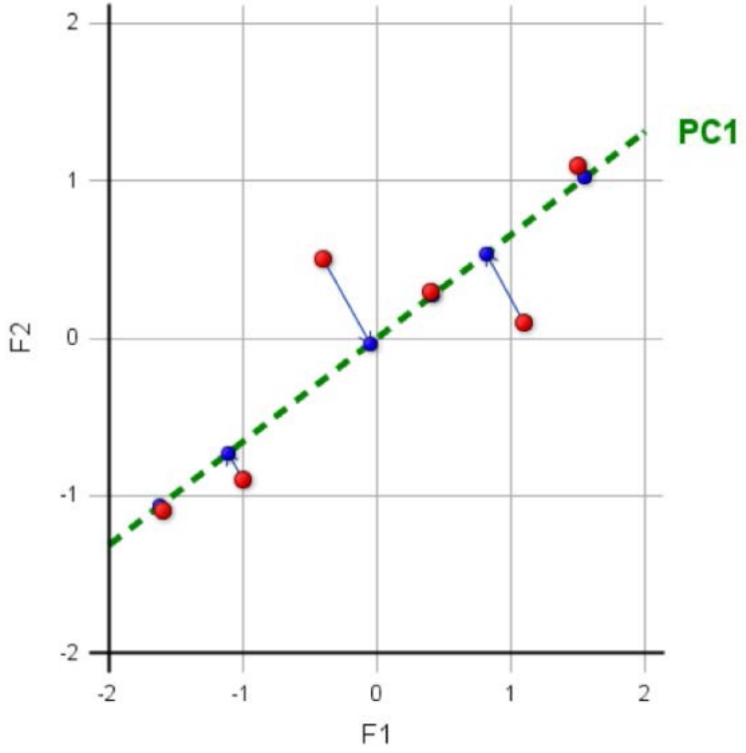
Principal Component Analysis



Feature	Variance
PC1	2.038
F1	1.468
F2	0.716

2.184

Principal Component Analysis

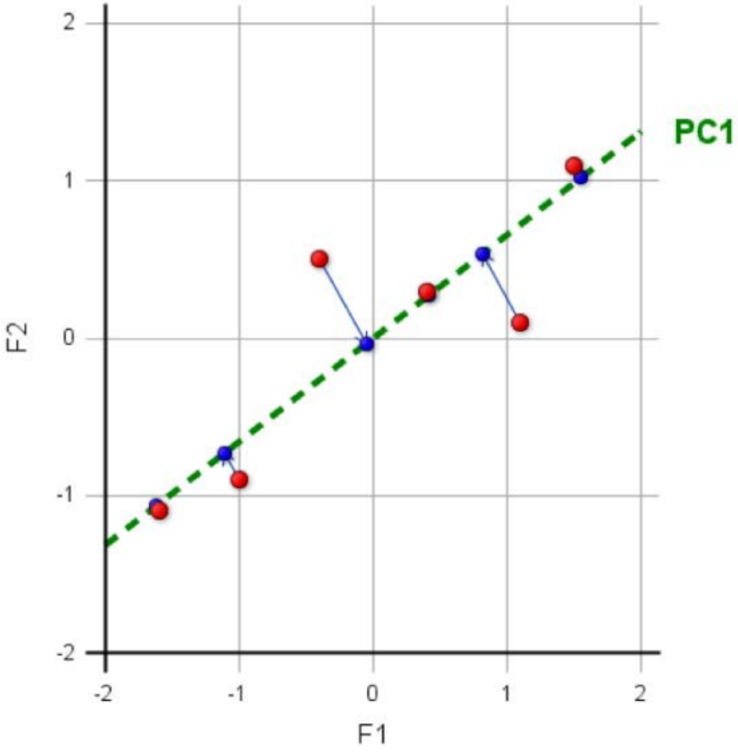


Feature	Variance
PC1	2.038
F1	1.468
F2	0.716

93%

2.184

Principal Component Analysis



Feature	Variance
PC1	2.038

Principal Component Analysis

$$\begin{bmatrix} 0.836 \\ 0.549 \end{bmatrix} \qquad \begin{bmatrix} -0.549 \\ 0.836 \end{bmatrix}$$

$$\lambda_1 = 2.038 \qquad \lambda_2 = 0.146$$

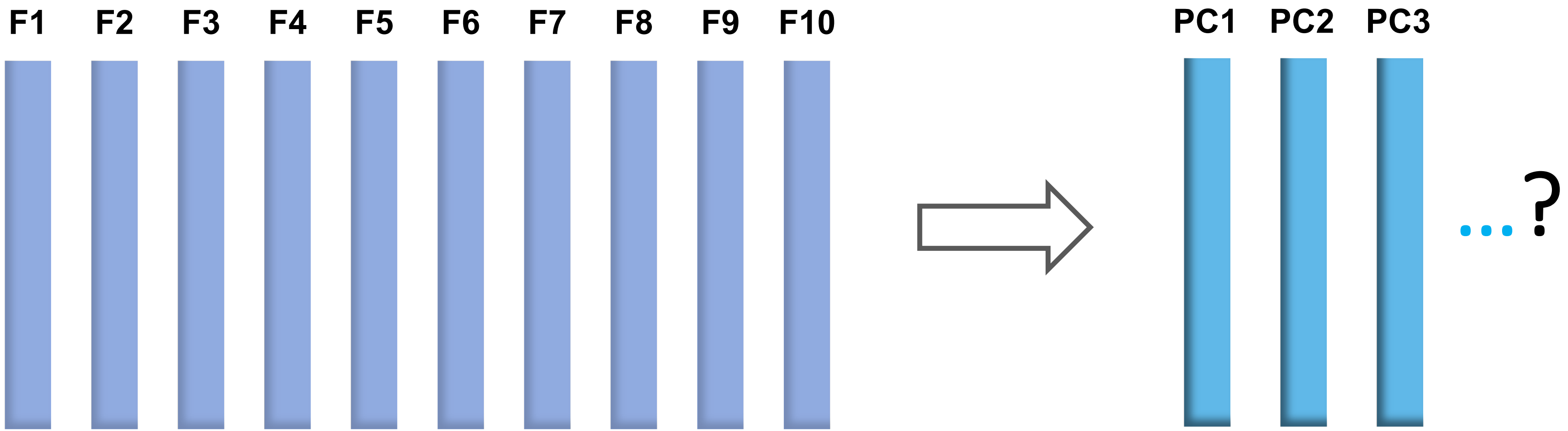
Principal Component Analysis

$$\begin{bmatrix} 0.836 \\ 0.549 \end{bmatrix}$$

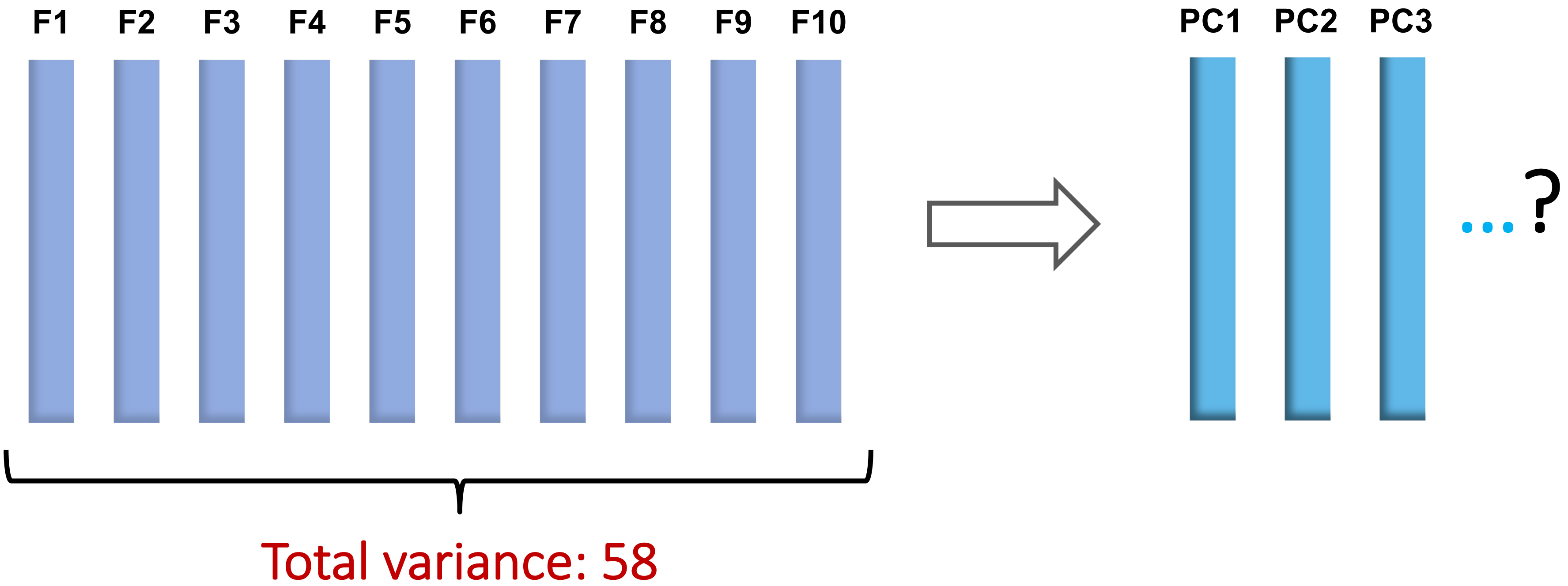
$\lambda_1 = 2.038$

Feature	Variance
PC1	2.038

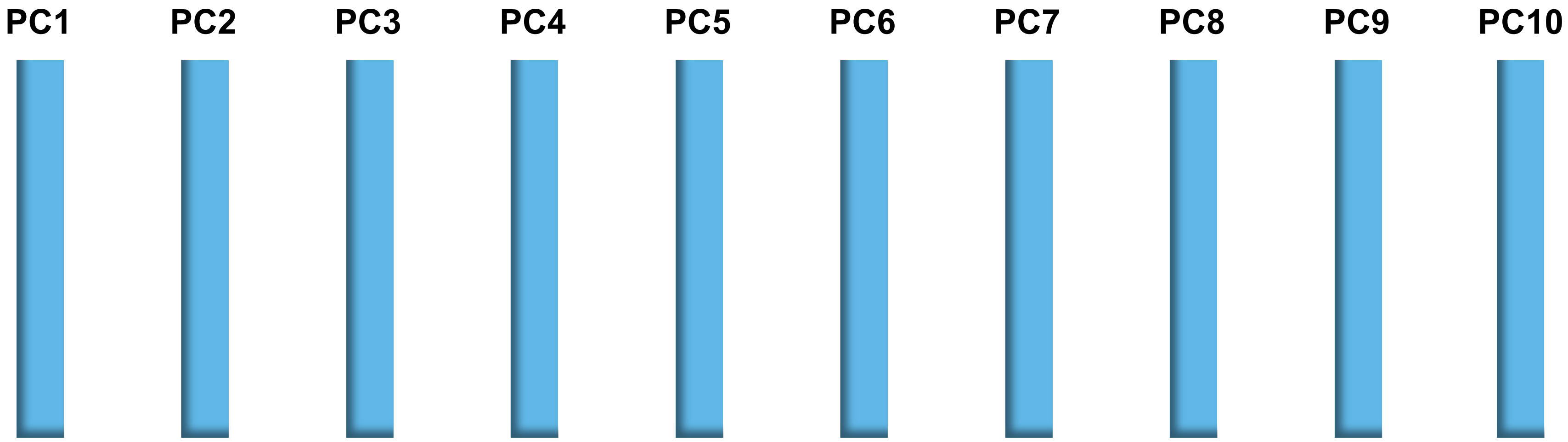
PCA in Higher Dimensions



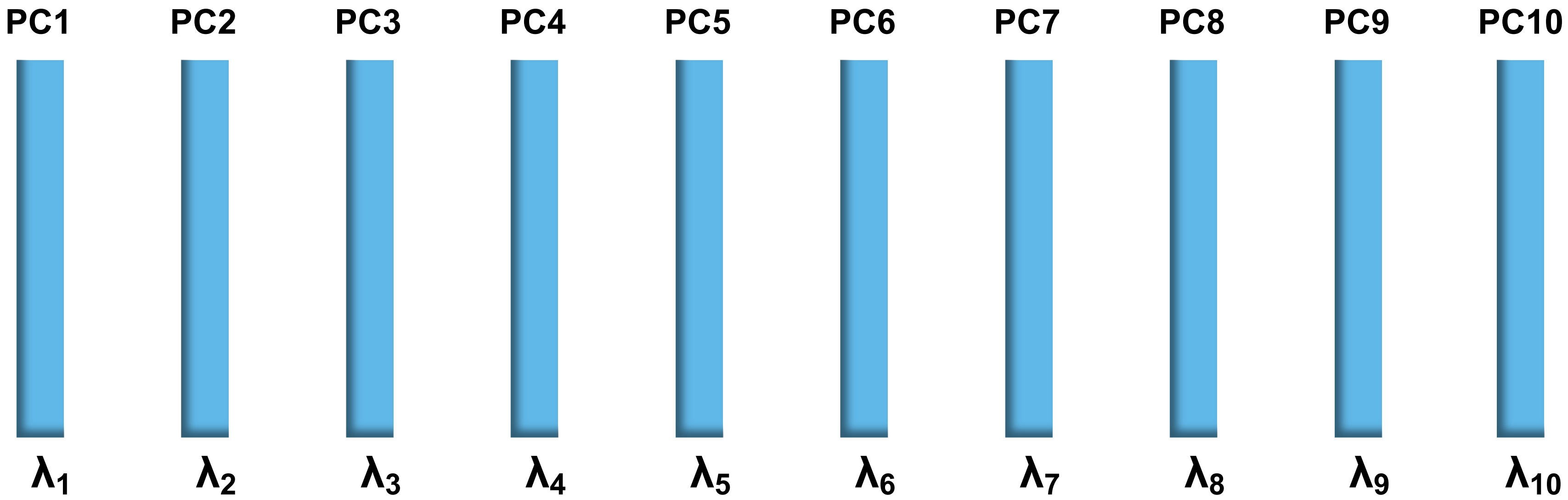
PCA in Higher Dimensions



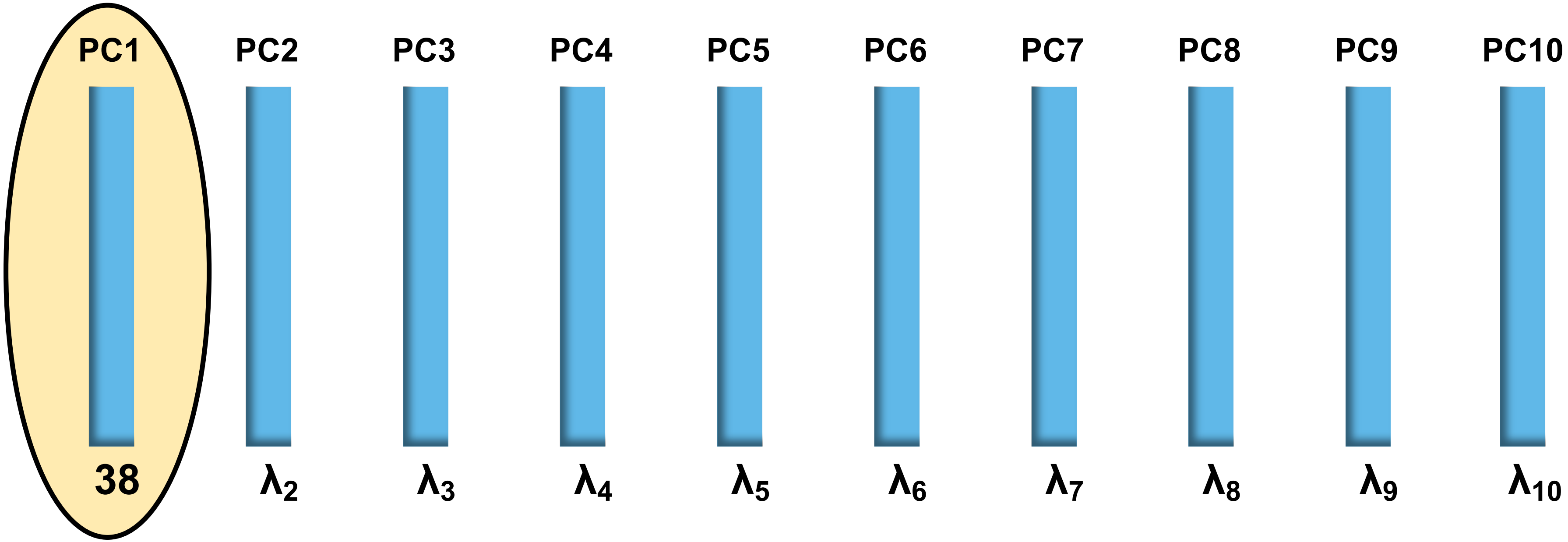
PCA in Higher Dimensions



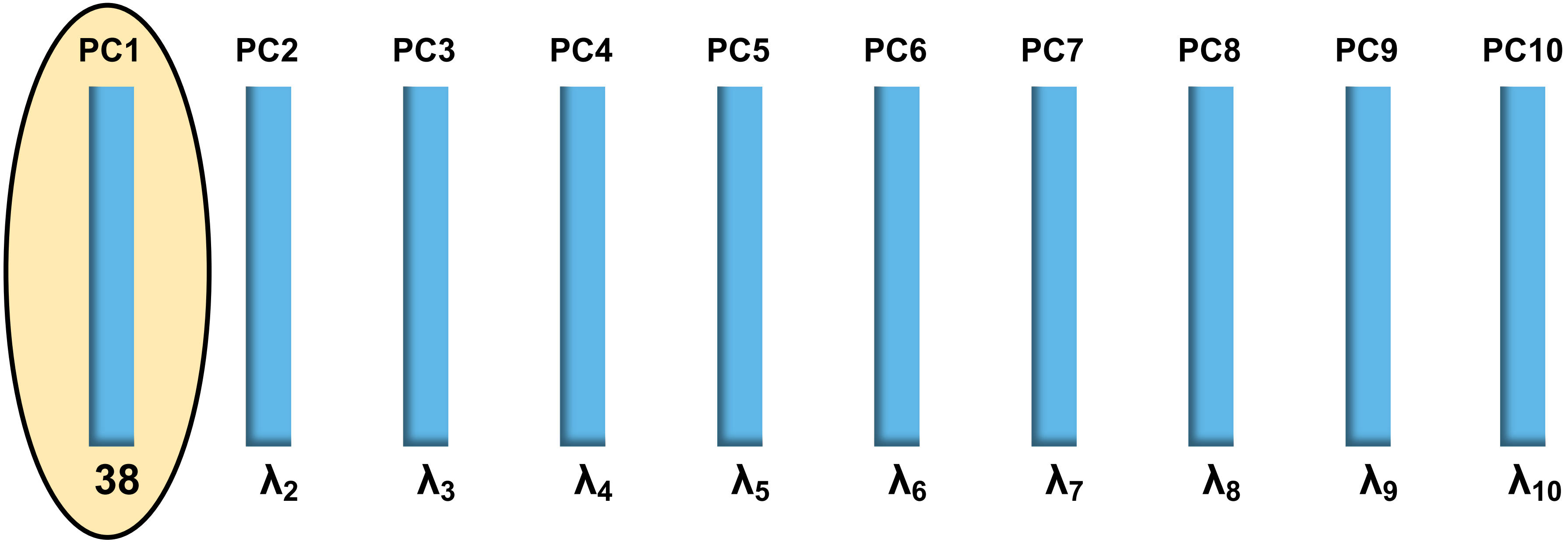
PCA in Higher Dimensions



PCA in Higher Dimensions

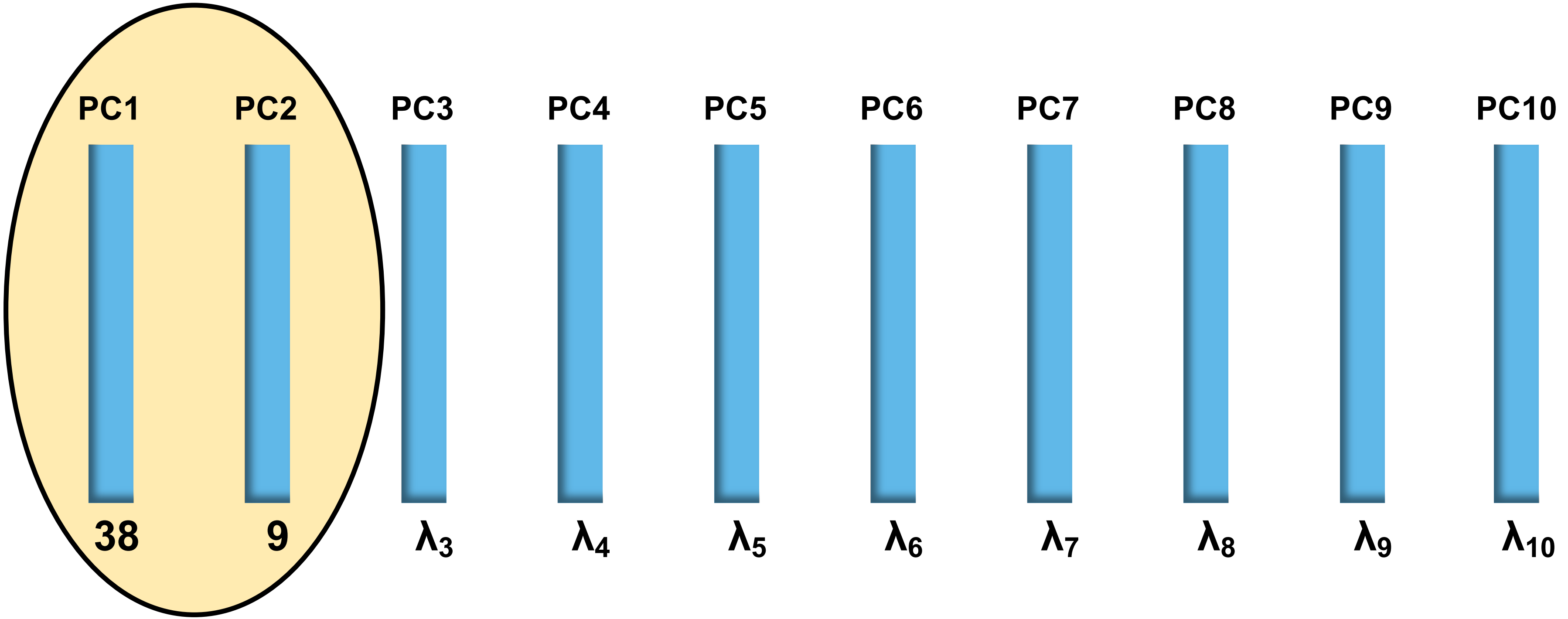


PCA in Higher Dimensions

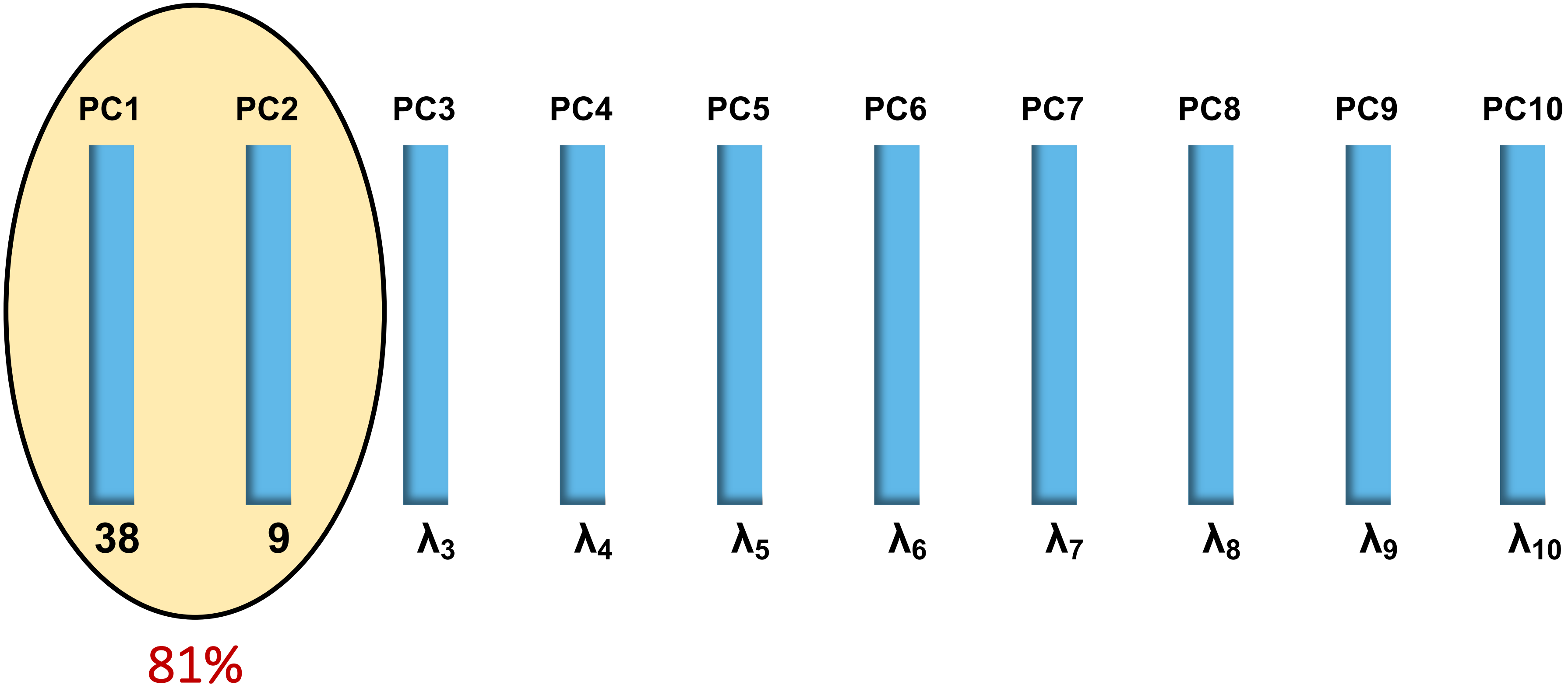


66%

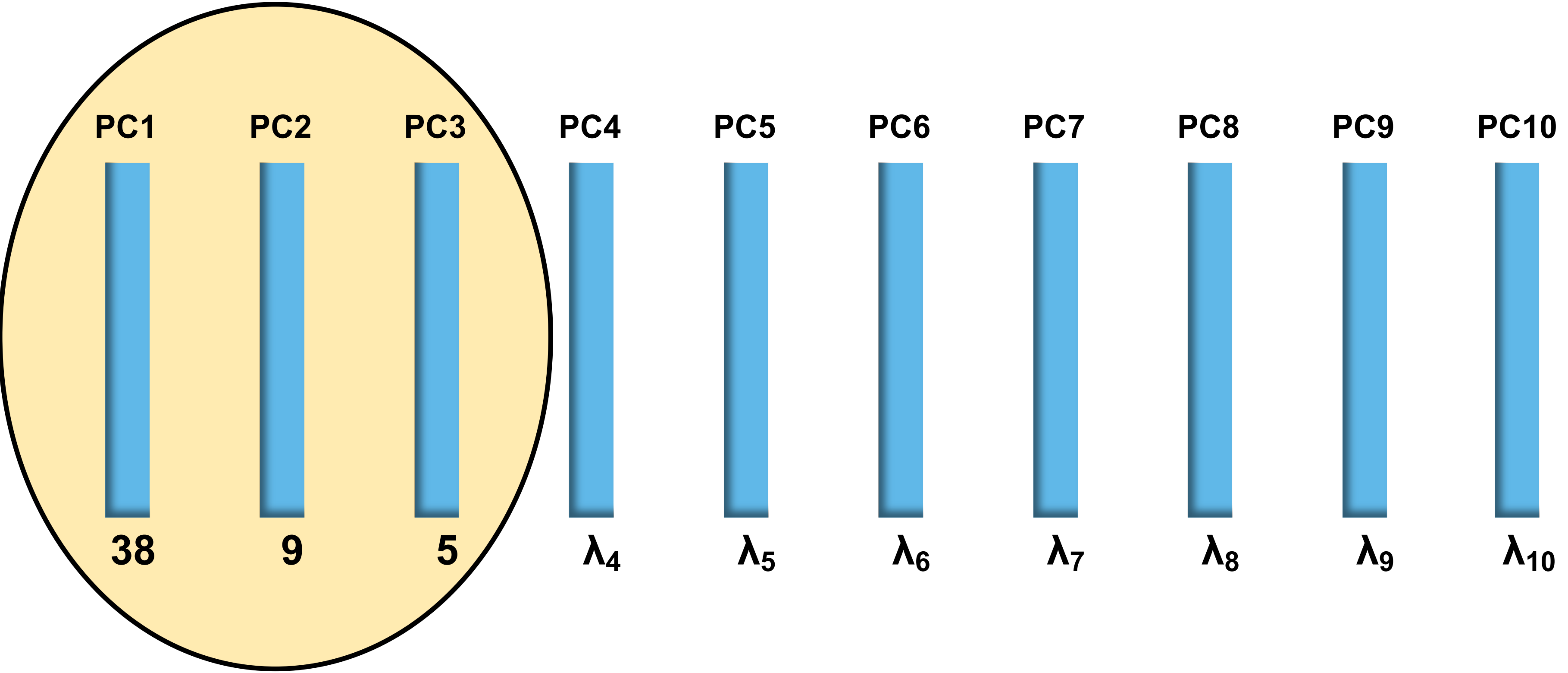
PCA in Higher Dimensions



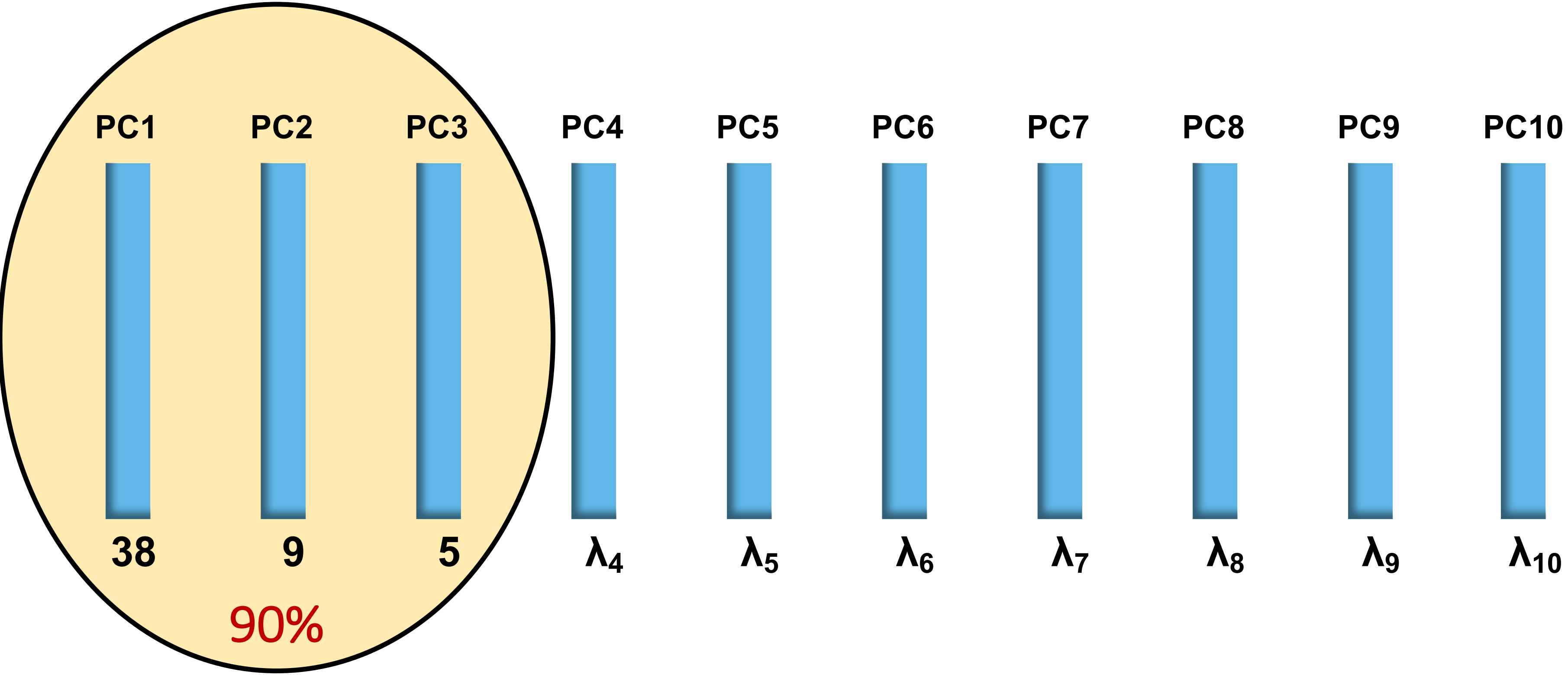
PCA in Higher Dimensions



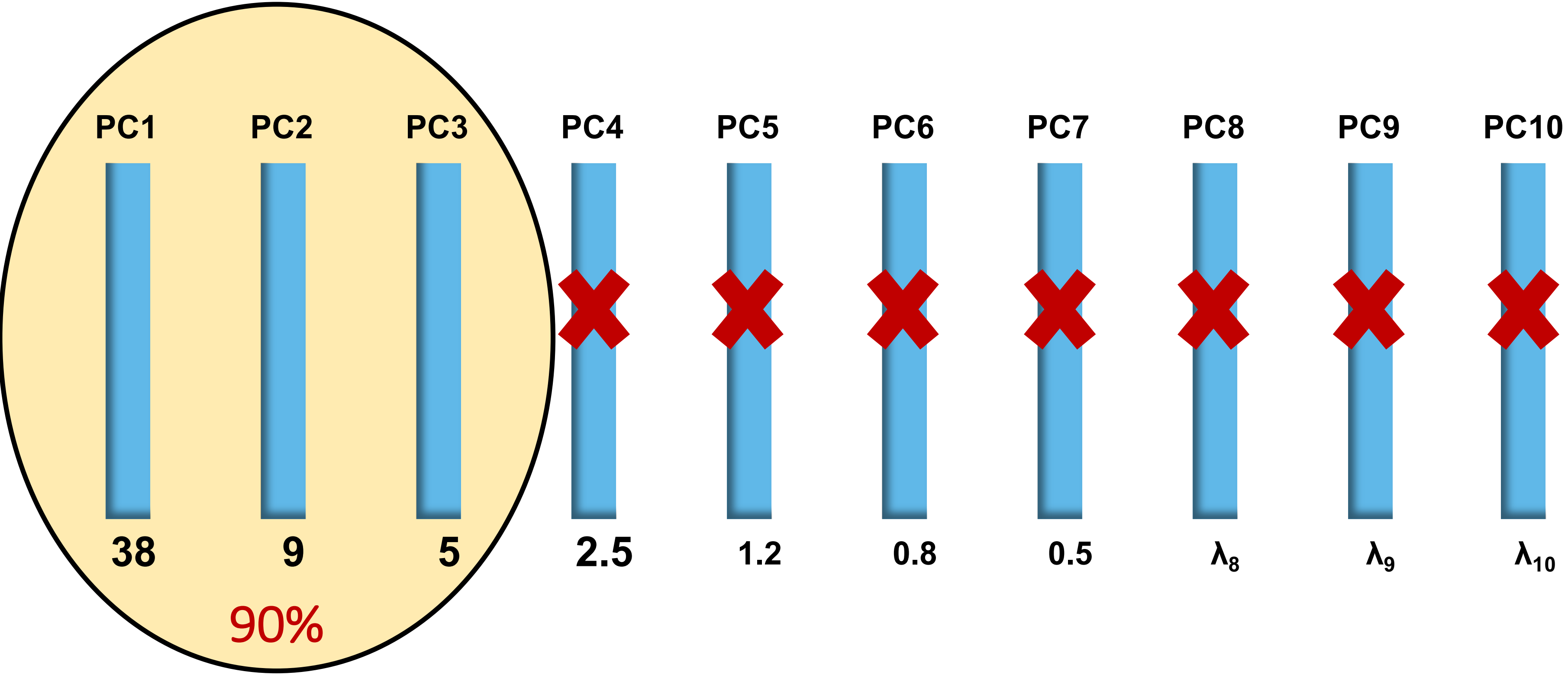
PCA in Higher Dimensions



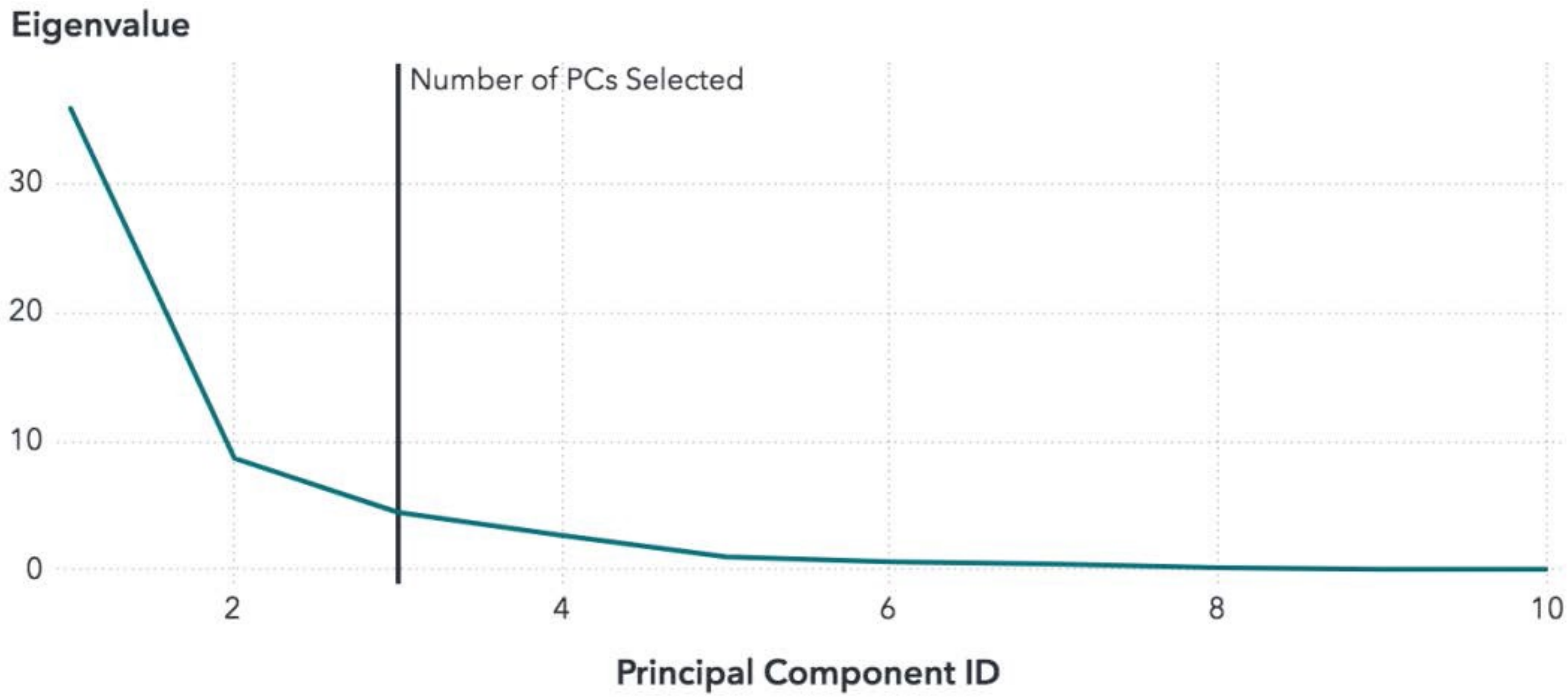
PCA in Higher Dimensions



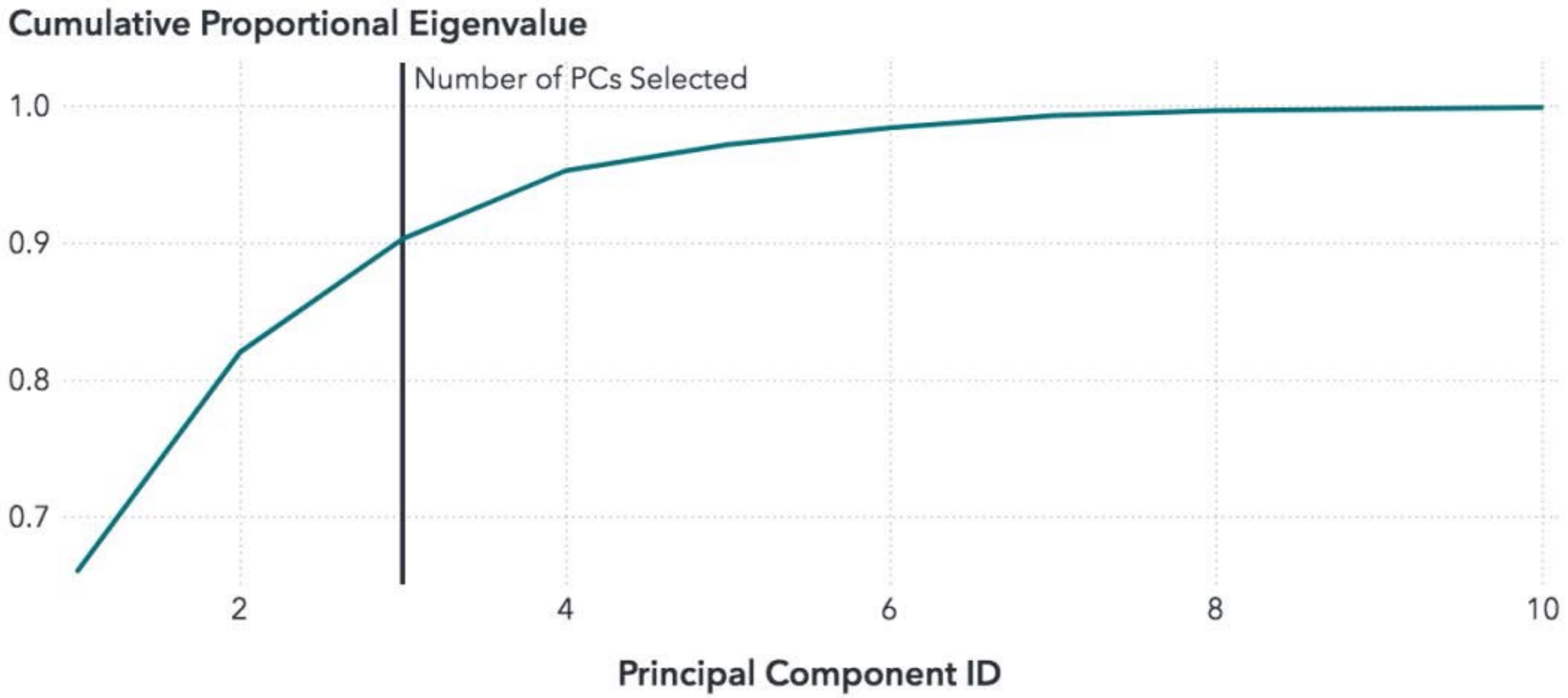
PCA in Higher Dimensions



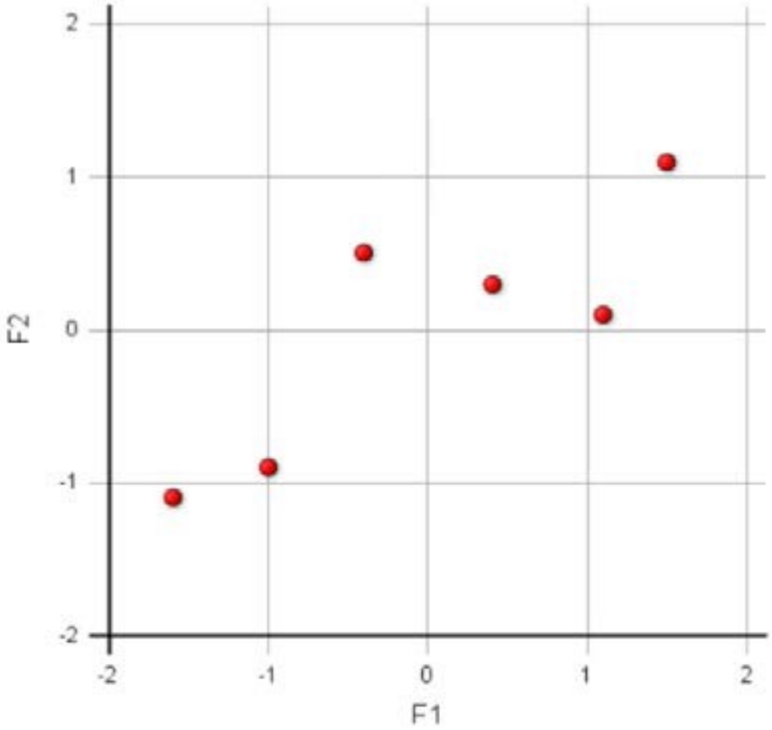
PCA in Higher Dimensions



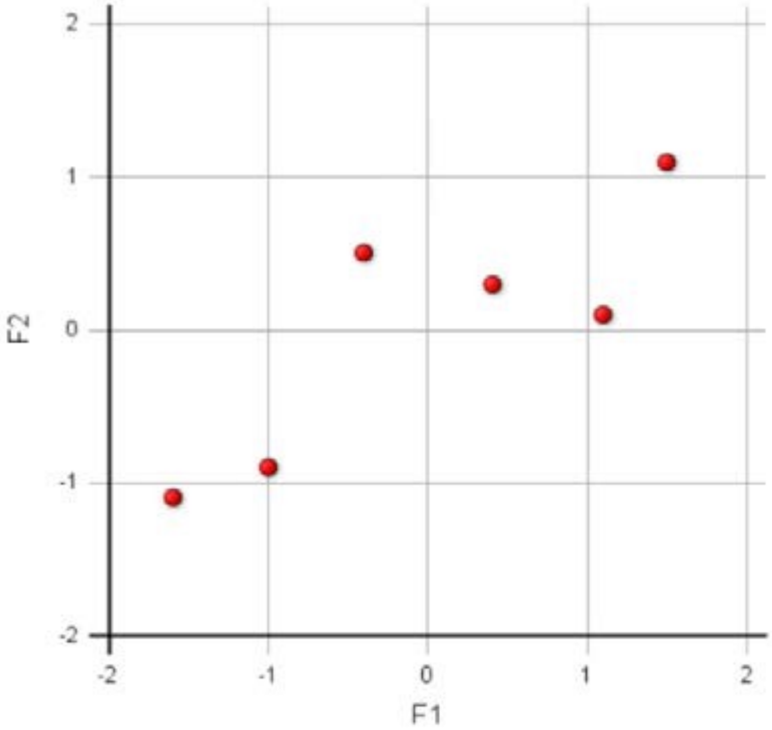
Principal Component Analysis

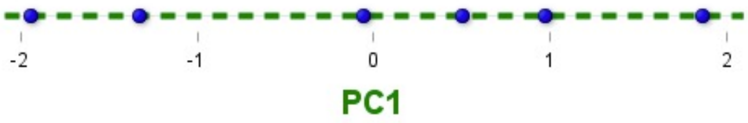


Reconstruction Error

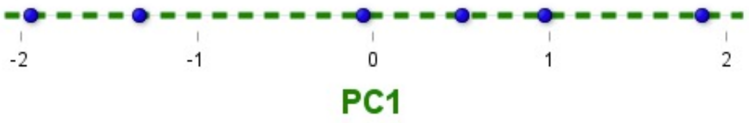


Reconstruction Error

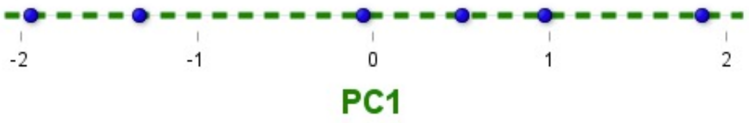


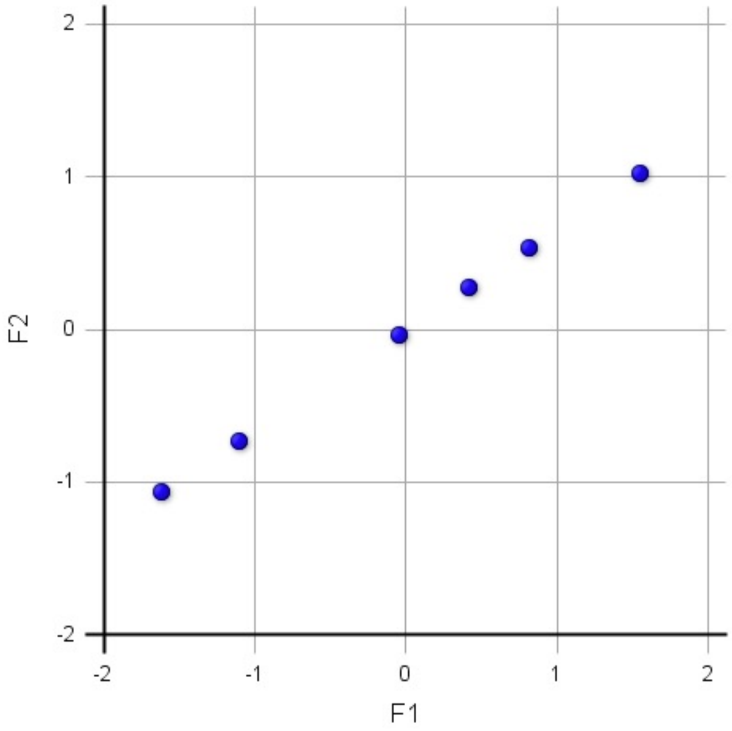


Reconstruction Error

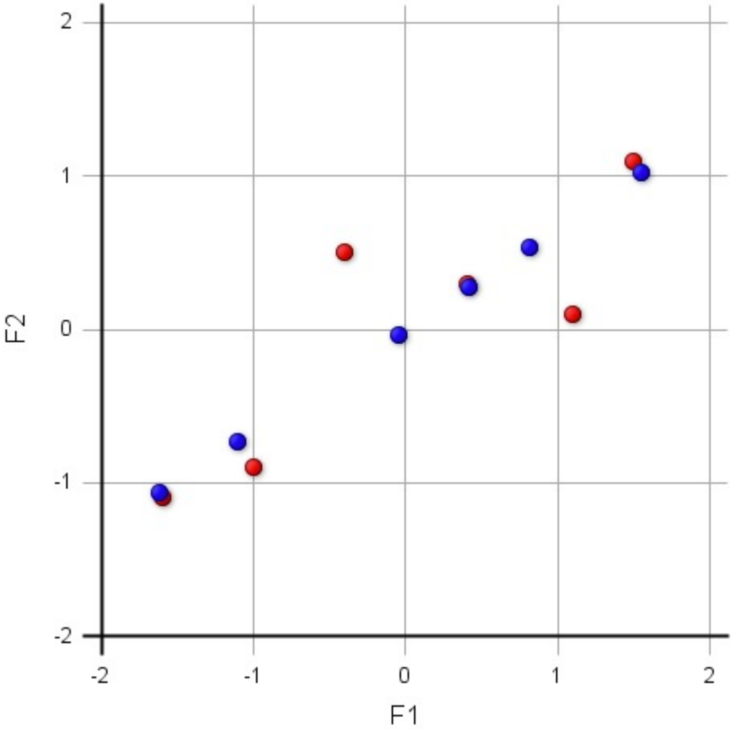


Reconstruction Error

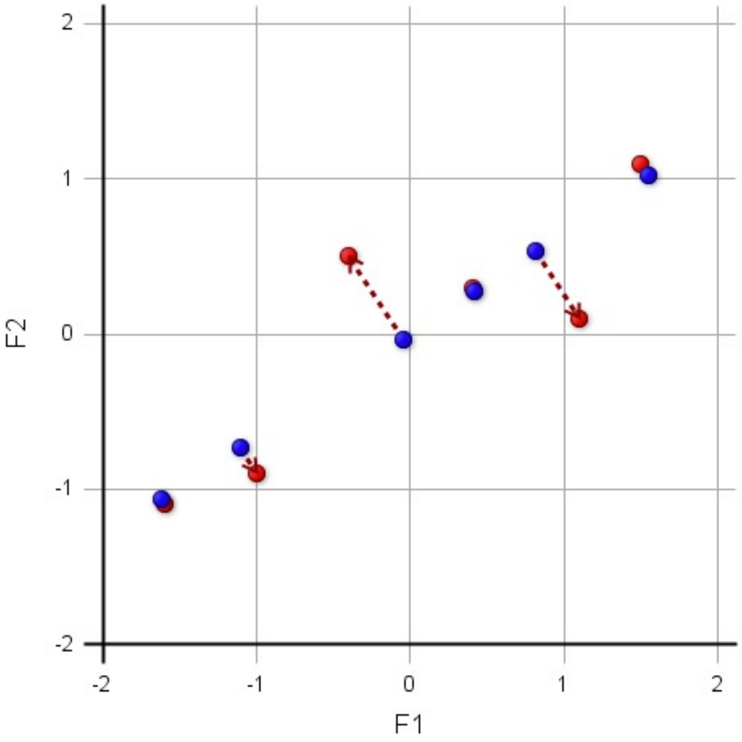




Reconstruction Error



Reconstruction Error



Summary

PCA is a dimensionality reduction technique.

Summary

PCA is a dimensionality reduction technique.

PCA finds the new feature set which retains maximum variance.

Summary

PCA is a dimensionality reduction technique.

PCA finds the new feature set which retains maximum variance.

Summary

PCA is a dimensionality reduction technique.

PCA finds the new feature set which $\underbrace{\text{retains maximum variance}}_{\text{minimizes reconstruction error}}$.

Summary

PCA is a dimensionality reduction technique.

PCA finds the new feature set which retains maximum variance.

minimizes reconstruction error

- New features are linear transformations of original features

Summary

PCA is a dimensionality reduction technique.

PCA finds the new feature set which $\underbrace{\text{retains maximum variance}}_{\text{minimizes reconstruction error}}$.

- New features are linear transformations of original features
- New features are linearly uncorrelated

Further Reading

- Jolliffe, I.T. 2002. *Principal Component Analysis*, 2nd ed. New York, NY: Springer.

Further Reading

- Jolliffe, I.T. 2002. *Principal Component Analysis*, 2nd ed. New York, NY: Springer.

- Shlens, Jonathon. 2014 “A Tutorial on Principal Component Analysis” https://arxiv.org/pdf/1404.1100.pdf
- Wiskott, Laurenz. 2013 “Lecture Notes on Principal Component Analysis” http://cs233.stanford.edu/ReferencedPapers/LectureNotes-PCA.pdf

Further Reading

- Jolliffe, I.T. 2002. *Principal Component Analysis*, 2nd ed. New York, NY: Springer.
- Shlens, Jonathon. 2014 "A Tutorial on Principal Component Analysis" https://arxiv.org/pdf/1404.1100.pdf
- Wiskott, Laurenz. 2013 "Lecture Notes on Principal Component Analysis" http://cs233.stanford.edu/ReferencedPapers/LectureNotes-PCA.pdf
- "Principal Component Analysis Explained Visually" by Victor Powell and Lewis Lehe http://setosa.io/ev/principal-component-analysis/

Thank you!

Caroline Walker
cwalker@warrenrogers.com