PSY9511: Seminar 5

Unsupervised learning

Esten H. Leonardsen

24.10.24

UNIVERSITY
OF OSLO

Outline

1. Overview of unsupervised learning
2. Clustering
    - K-means
    - Hierarchical
3. Dimensionality reduction
    - Principal component analysis (PCA)
    - Independent component analysis (ICA)
    - Partial least squares (PLS)

Unsupervised learning

UNIVERSITY
OF OSLO

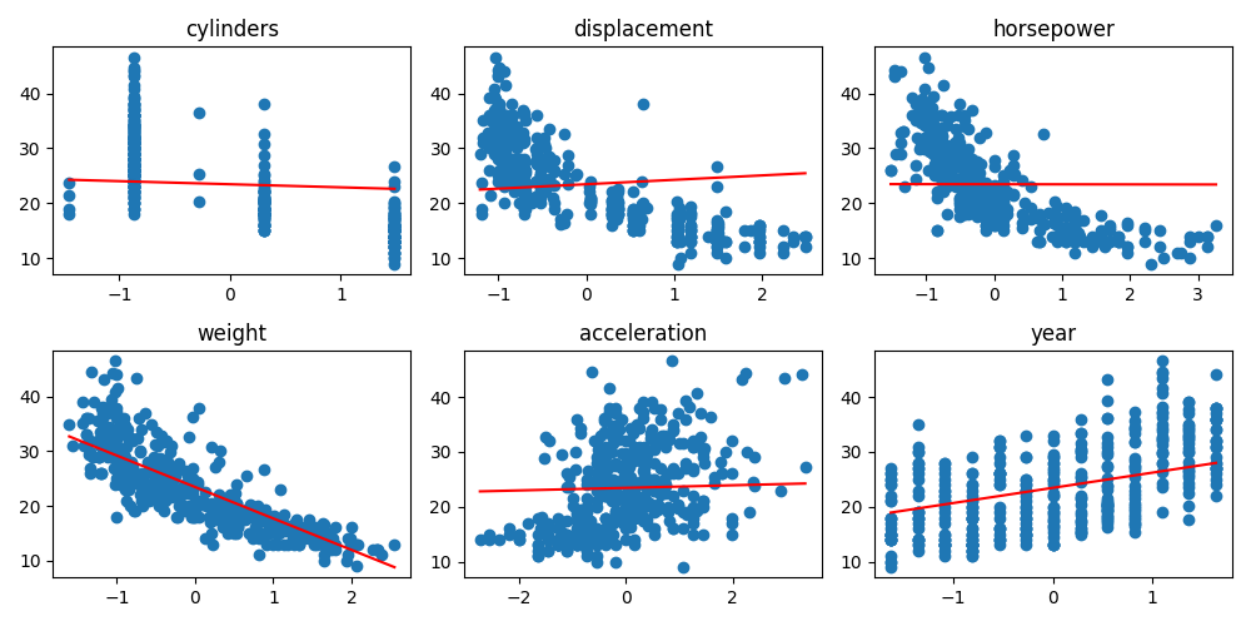
Unsupervised learning: Motivation

Supervised learning: Find $\hat{y} = f(X)$

Unsupervised learning: Motivation

Supervised learning: Find $\hat{y} = f(X)$

- Descriptive: Understand the relationship between X and y



Unsupervised learning: Motivation

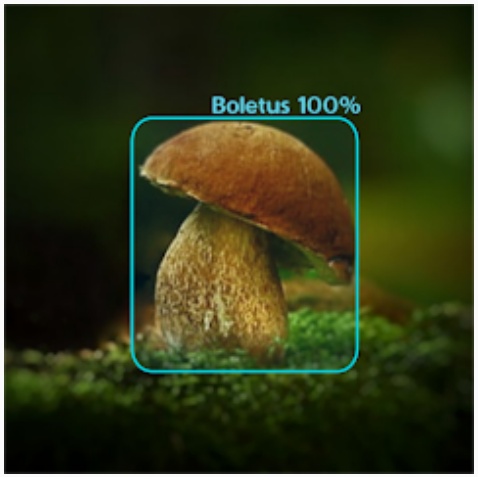
Supervised learning: Find $\hat{y} = f(X)$

- Descriptive: Understand the relationship between X and y
- Predictive: Predict y given new X.

Unsupervised learning: Motivation

Supervised learning: Find $\hat{y} = f(X)$

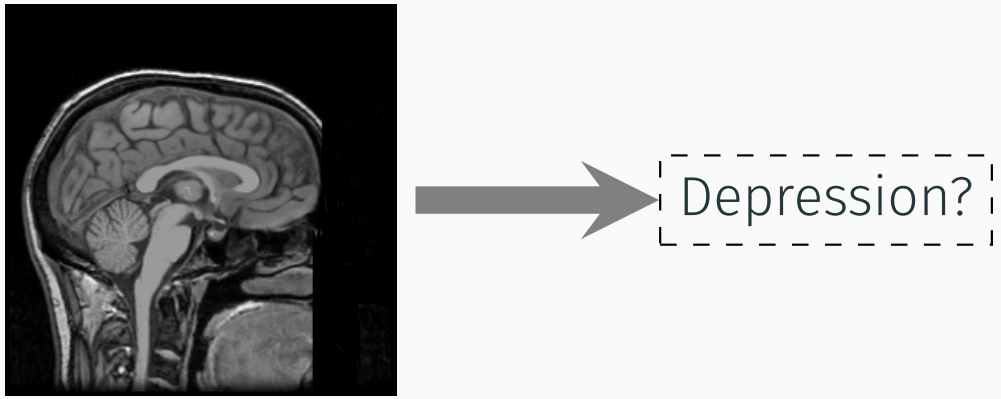
- Descriptive: Understand the relationship between X and y
- Predictive: Predict y given new X.
  - Because the predictions are useful



Unsupervised learning: Motivation

Supervised learning: Find $\hat{y} = f(X)$

- Descriptive: Understand the relationship between X and y
- Predictive: Predict y given new X.
  - Because the predictions are useful
  - Because we want to know if it is possible



Unsupervised learning: Motivation

Supervised learning: Find $\hat{y} = f(X)$

- Descriptive: Understand the relationship between X and y
- Predictive: Predict y given new X.
  - Because the predictions are useful
  - Because we want to know if it is possible

Unsupervised learning: Are there some interesting patterns in X?

Unsupervised learning: Motivation

Supervised learning: Find $\hat{y} = f(X)$

- Descriptive: Understand the relationship between X and y
- Predictive: Predict y given new X.
  - Because the predictions are useful
  - Because we want to know if it is possible

Unsupervised learning: Are there some interesting patterns in X?

- Can we find subgroups or interesting axes of variability?

Unsupervised learning: Motivation

Supervised learning: Find $\hat{y} = f(X)$

- Descriptive: Understand the relationship between X and y
- Predictive: Predict y given new X.
  - Because the predictions are useful
  - Because we want to know if it is possible

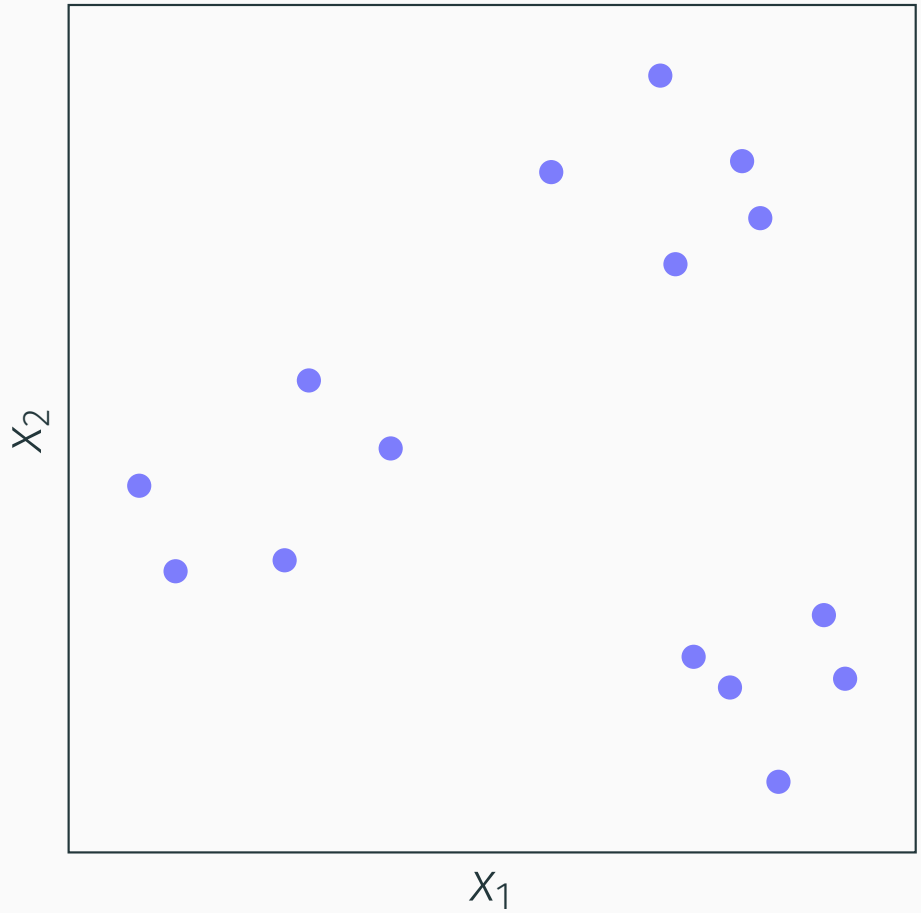
Unsupervised learning: Are there some interesting patterns in X?

- Can we find subgroups or interesting axes of variability?
- Visualization

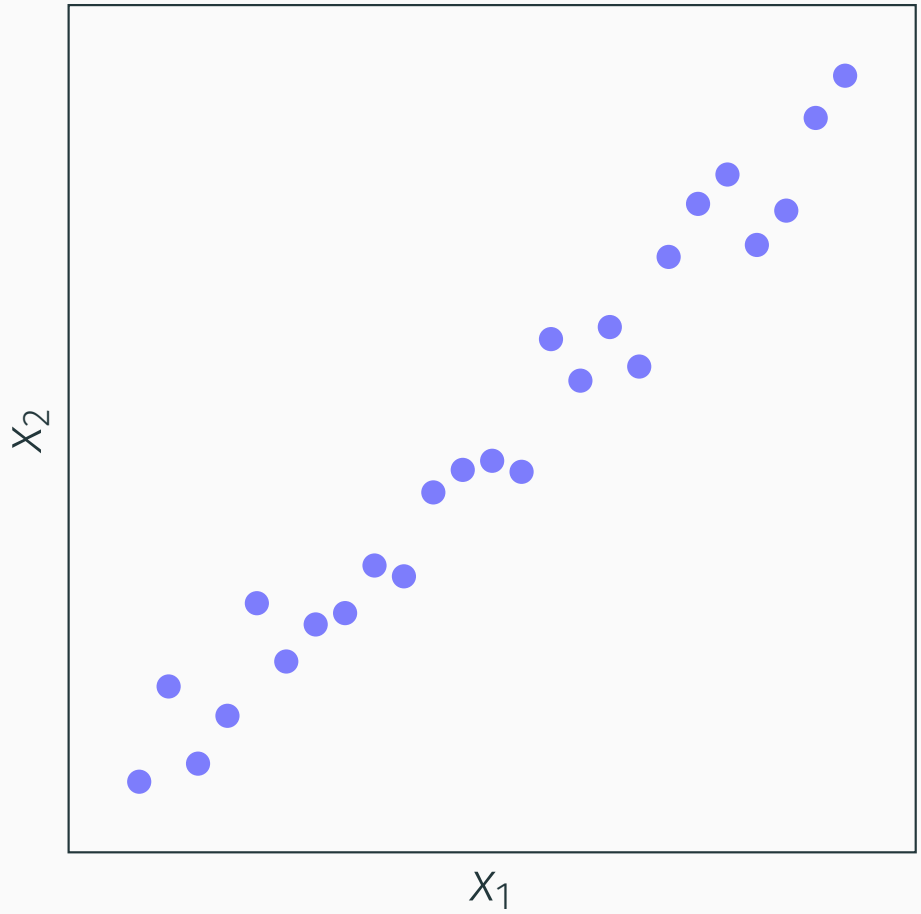
Dimensionality reduction

UNIVERSITY
OF OSLO

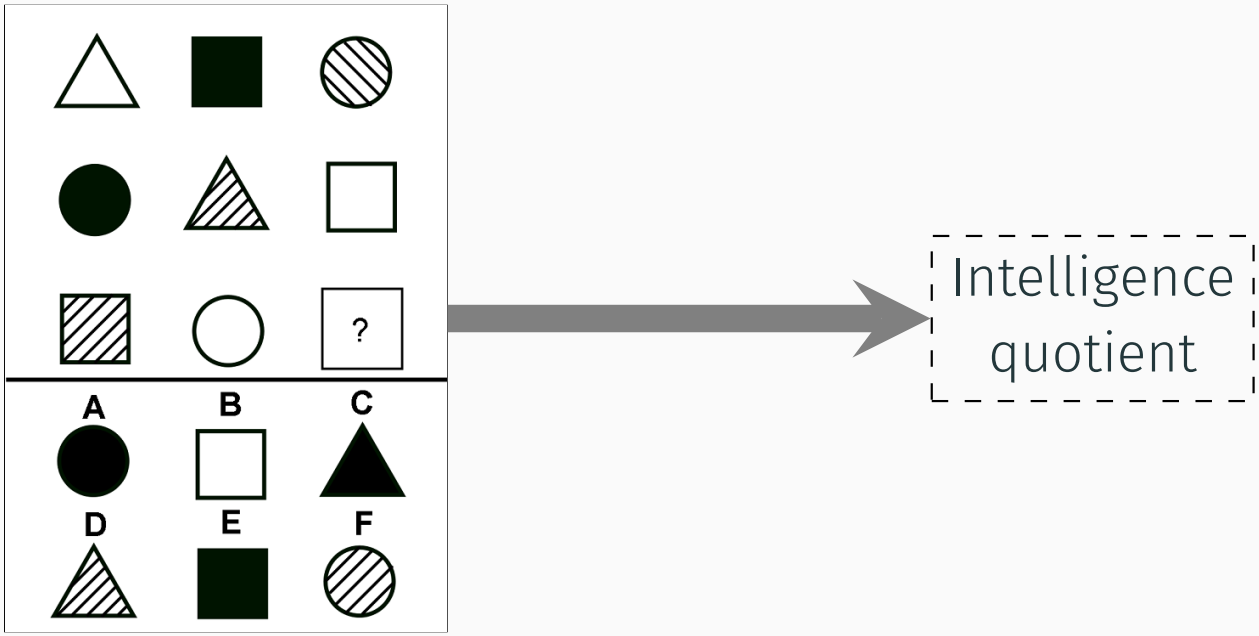
Dimensionality reduction: Motivation



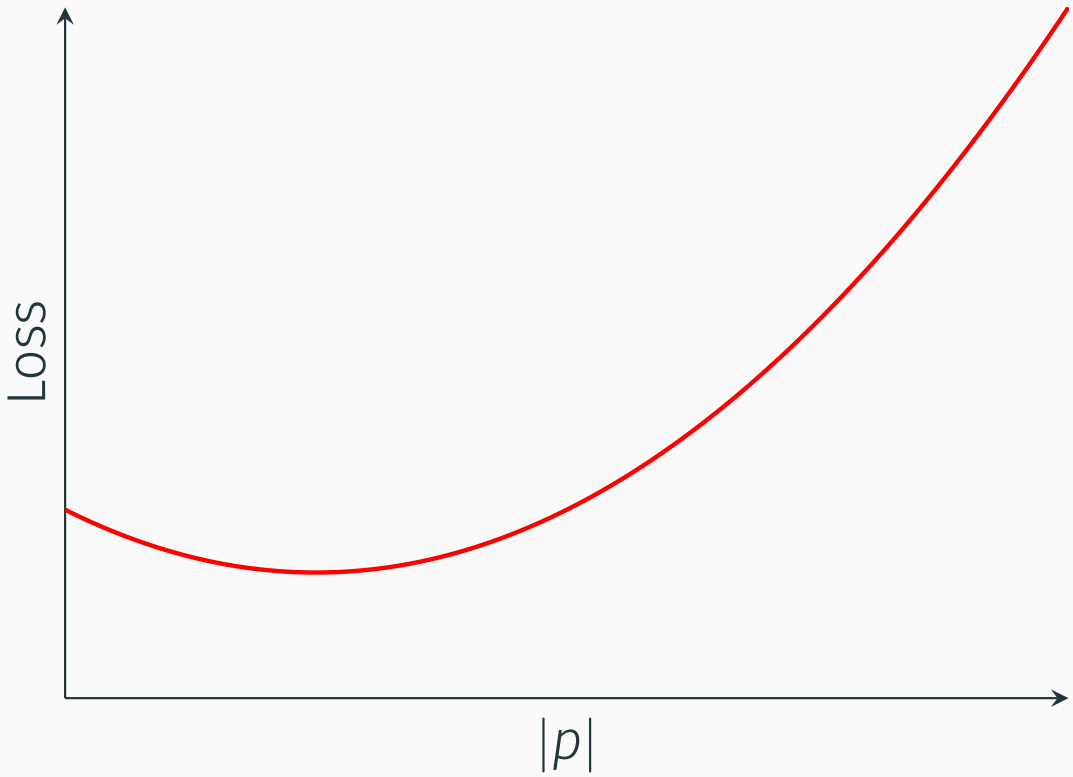
Dimensionality reduction: Motivation



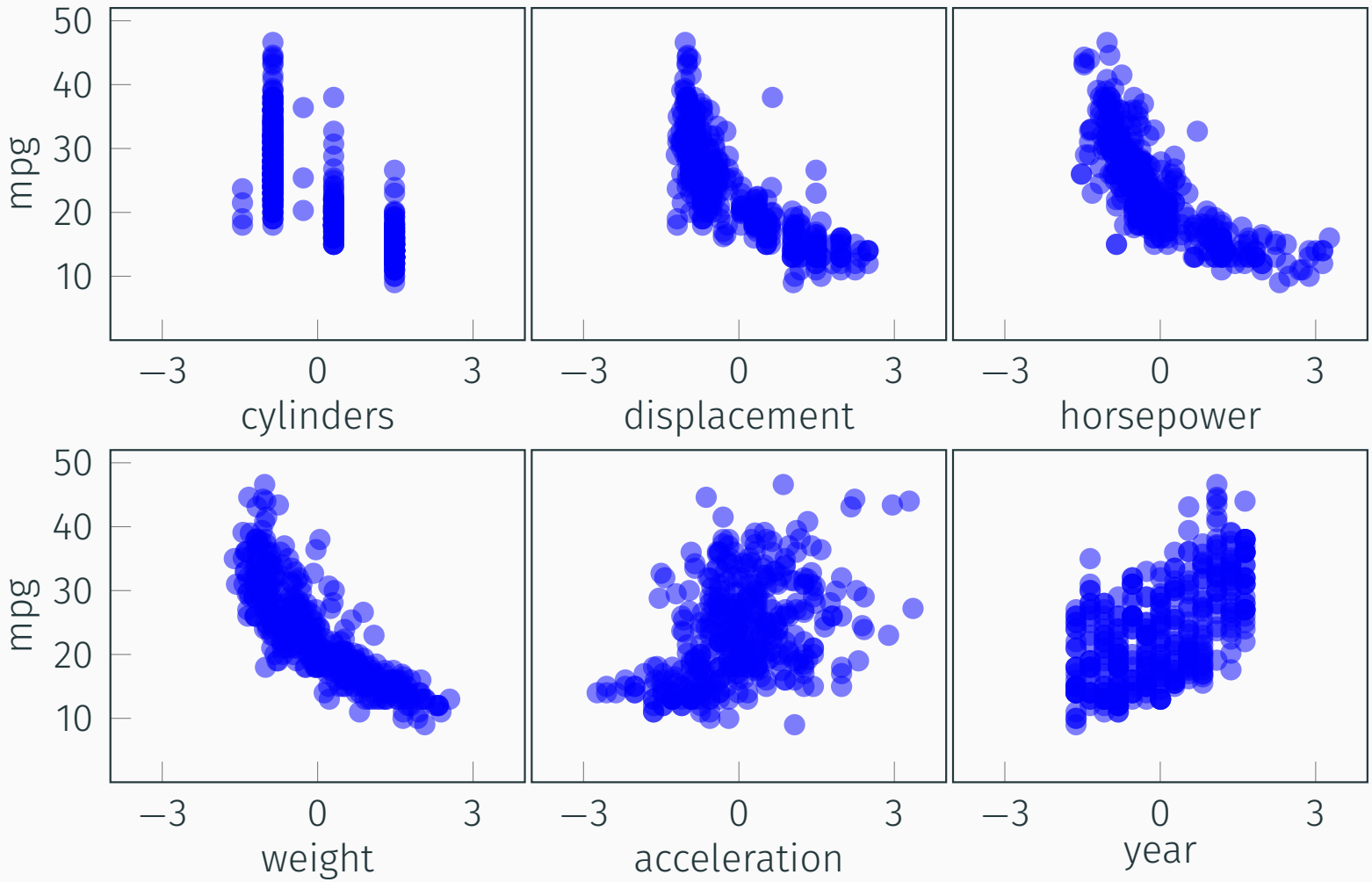
Dimensionality reduction: Motivation



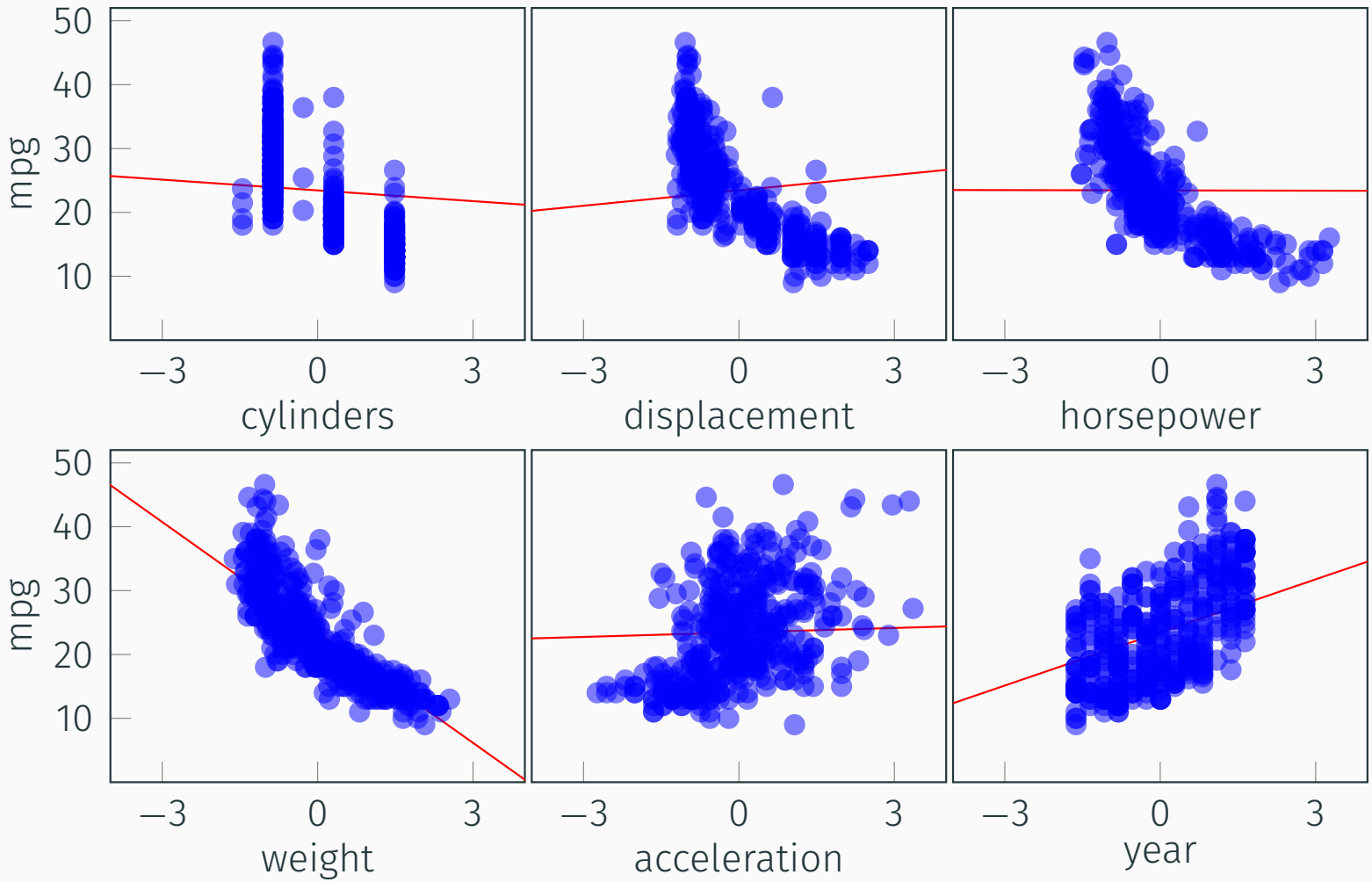
Dimensionality reduction: Motivation



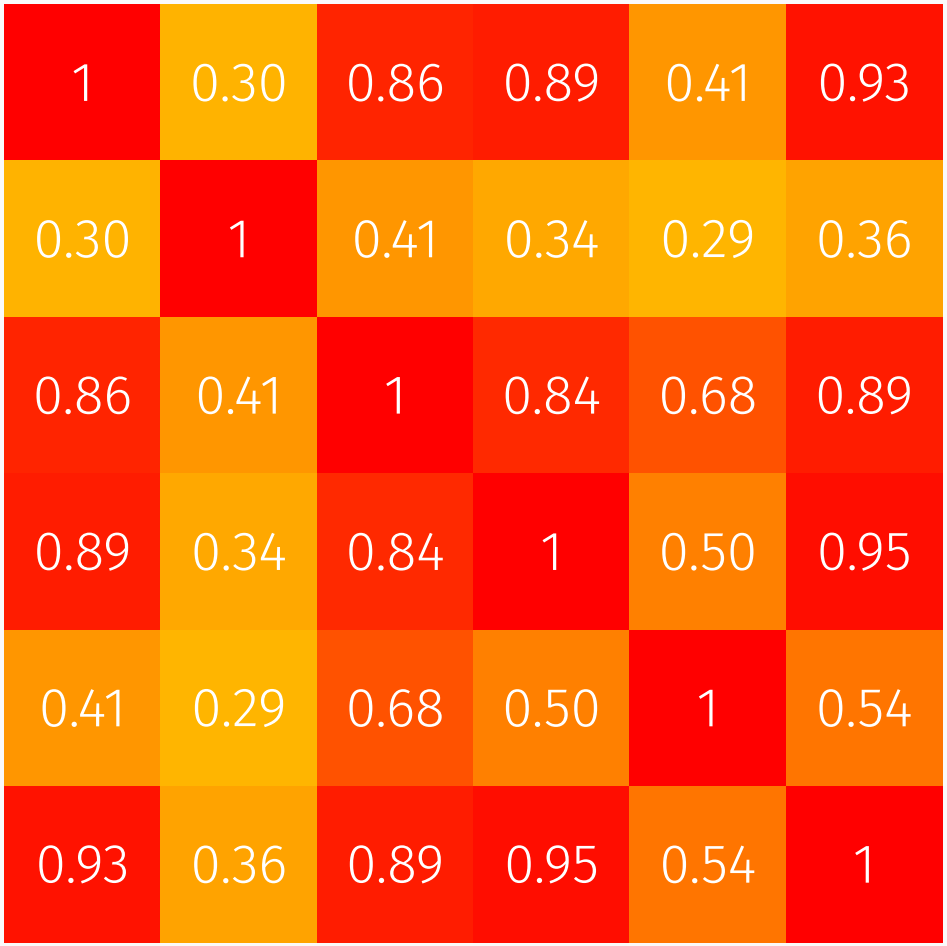
Dimensionality reduction: Motivation



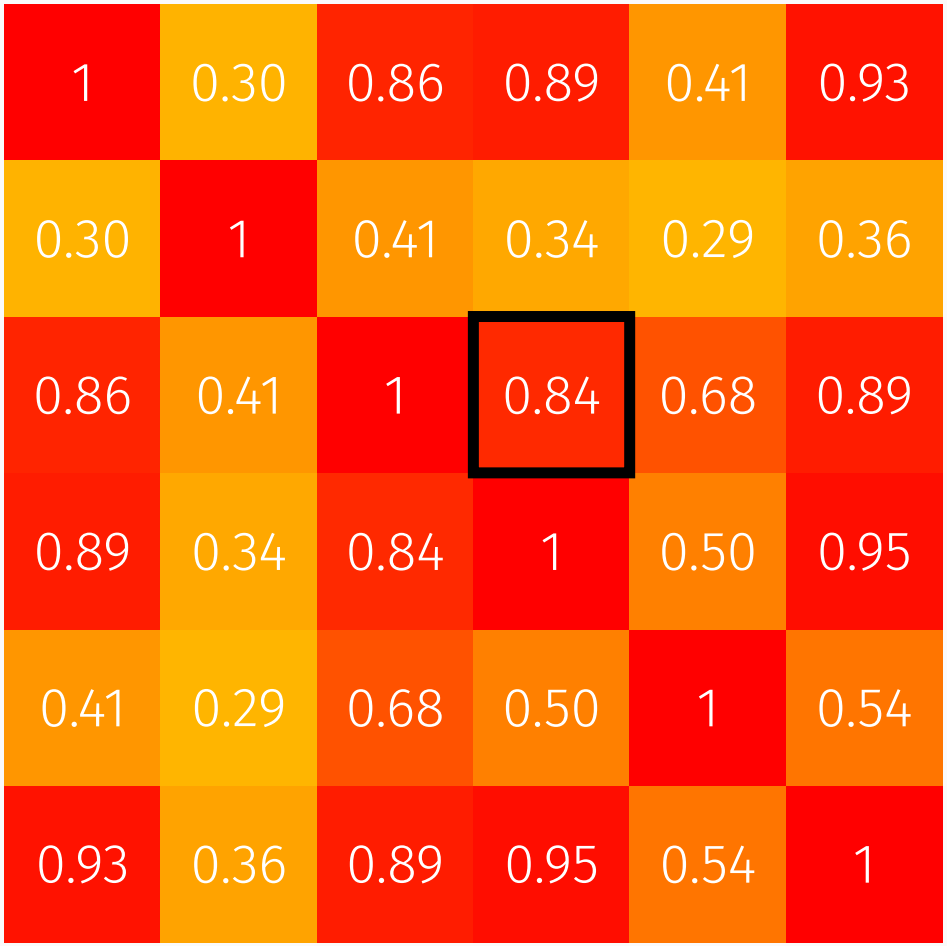
Dimensionality reduction: Motivation



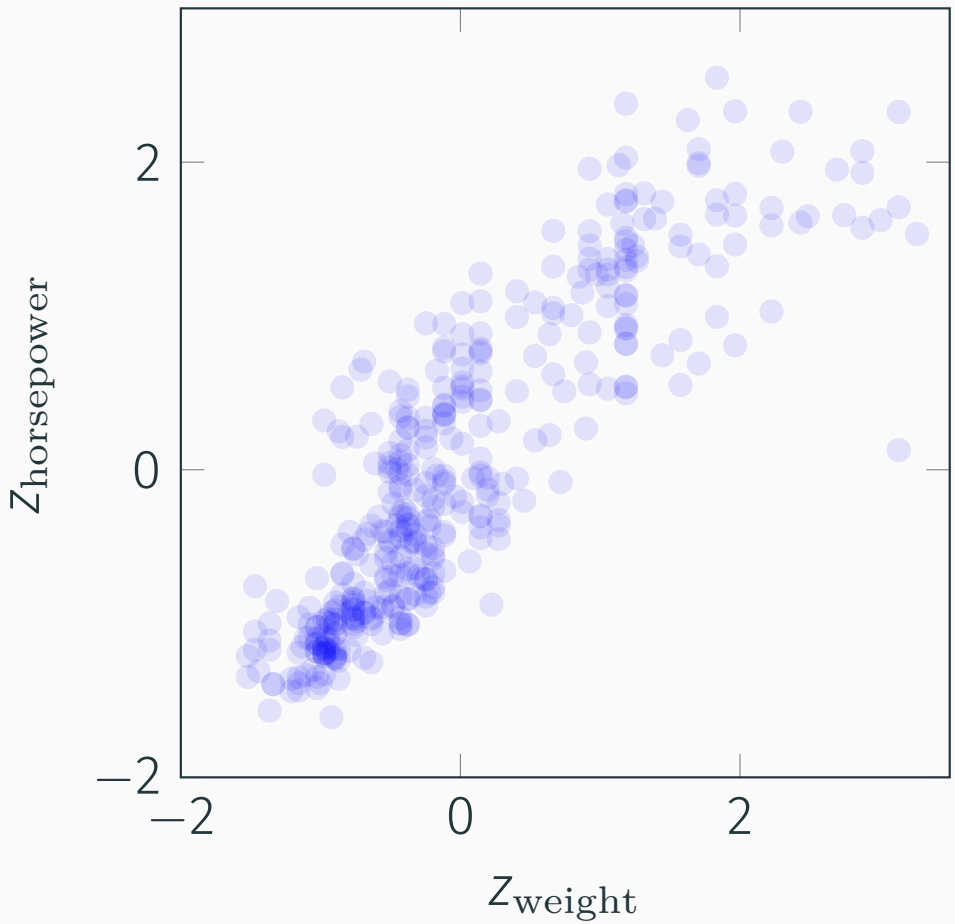
Dimensionality reduction: Motivation



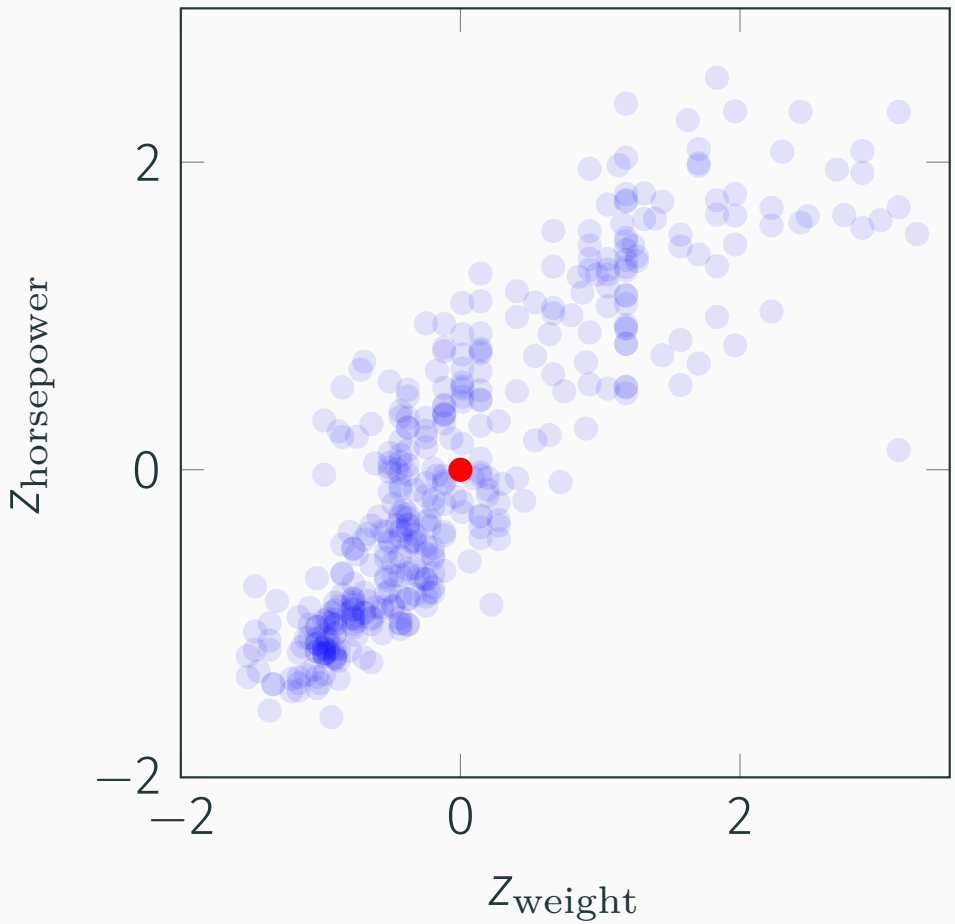
Dimensionality reduction: Motivation



Dimensionality reduction: Principal component analysis

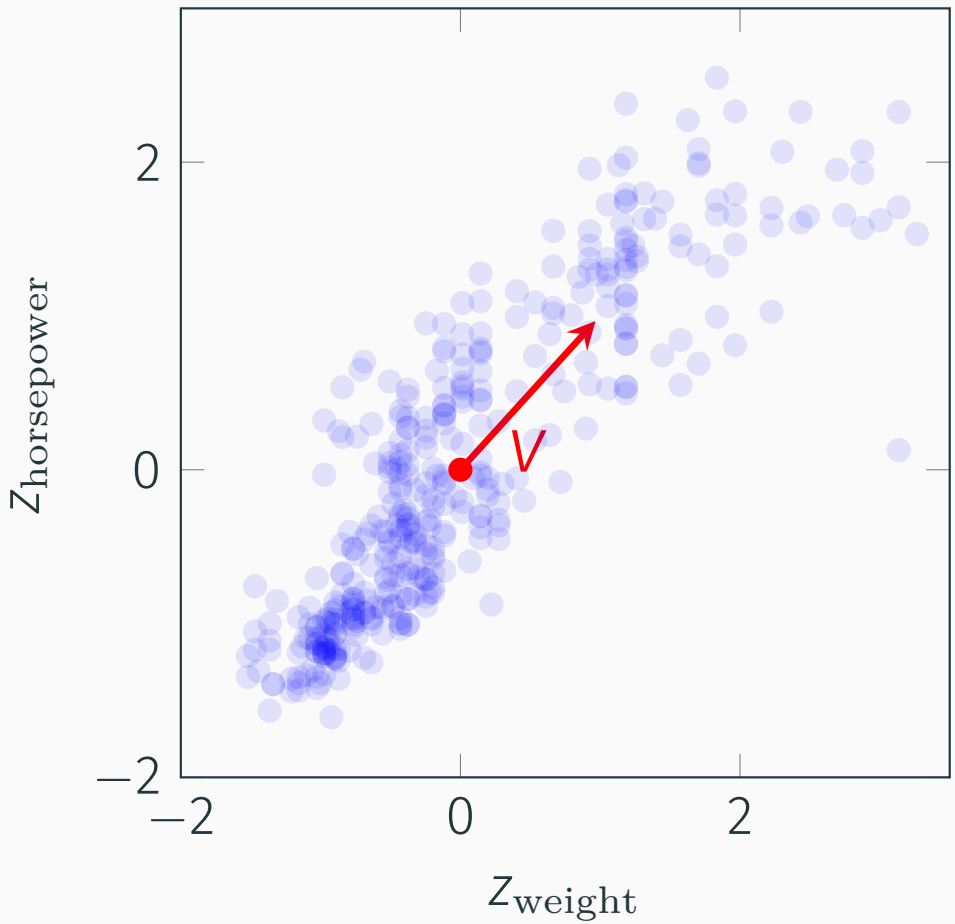


Dimensionality reduction: Principal component analysis



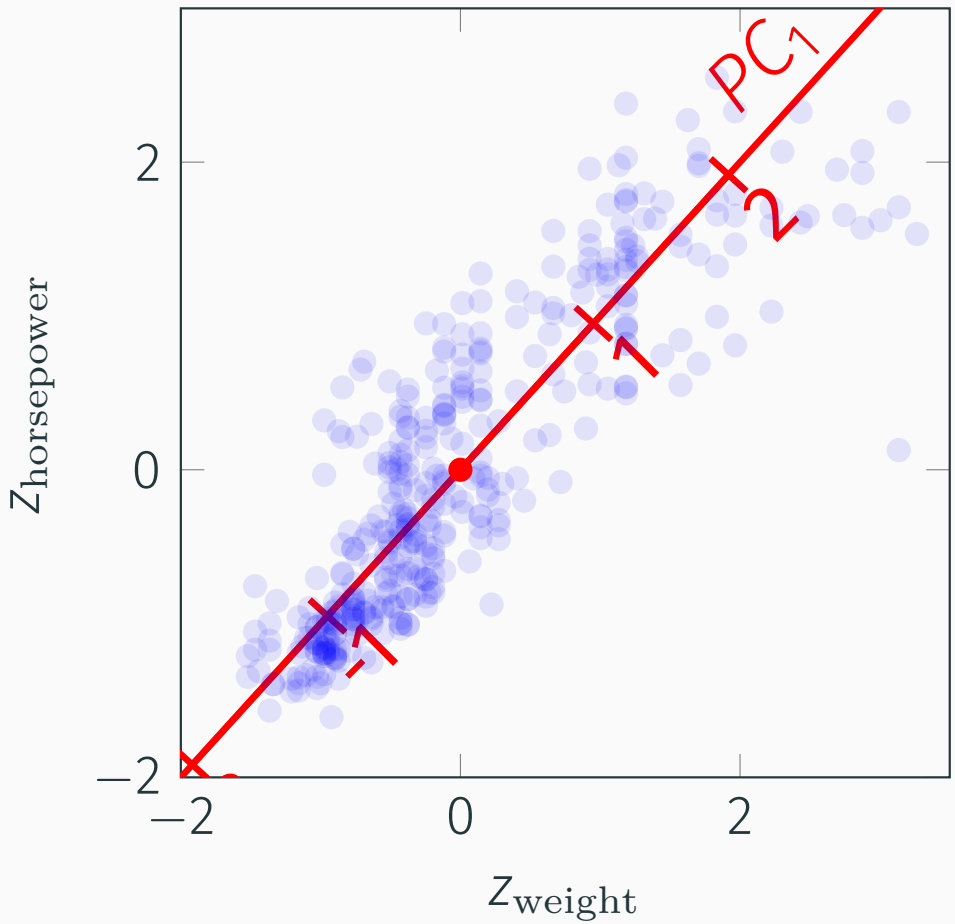
$c \rightarrow$ center of the data

Dimensionality reduction: Principal component analysis



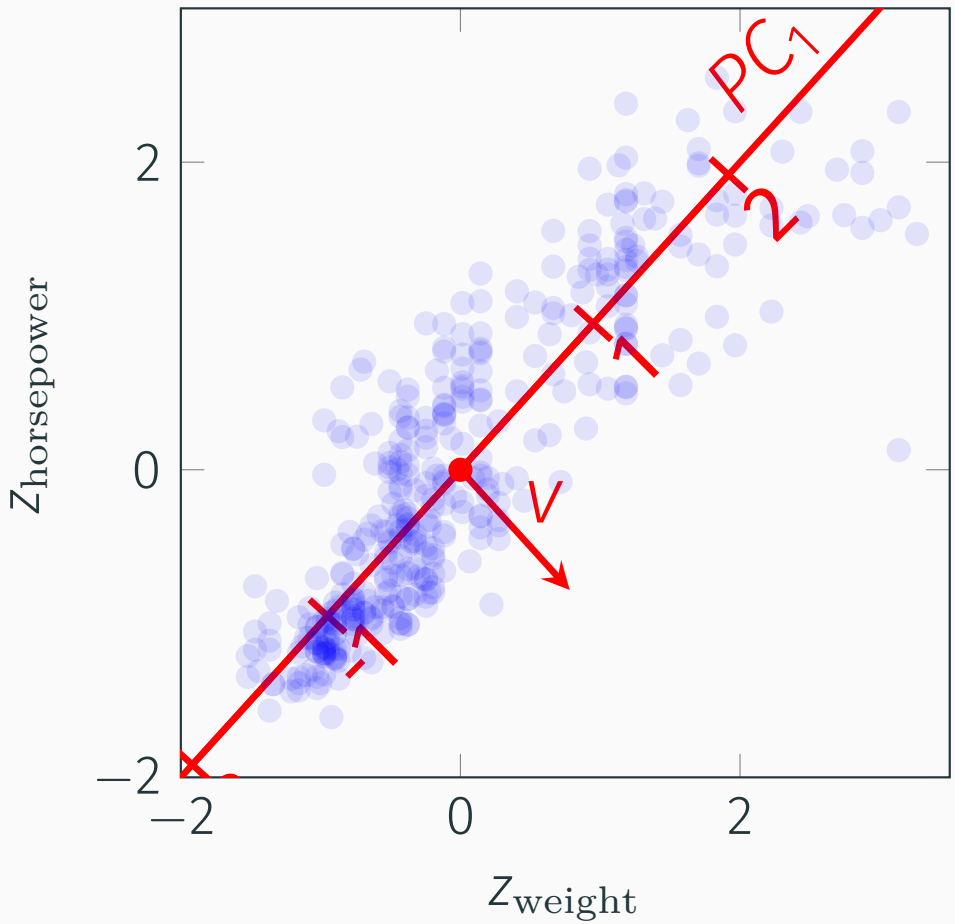
$v \rightarrow$ direction of maximum variance

Dimensionality reduction: Principal component analysis



$$PC_1 \rightarrow 0.69 * z_{\text{horsepower}} + 0.71 * z_{\text{weight}}$$

Dimensionality reduction: Principal component analysis



$v \rightarrow$ direction of maximum variance **orthogonal** to PC_1