1 Prediction with SVM

Recall our formulation of the SVM dual:

$$\max_{\alpha} \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y^{(i)} y^{(j)} x^{(i)T} x^{(j)}$$

such that

$$\alpha_i \geq 0, i = 1, ..., N$$

Suppose we get a new point x and we would like to predict its label (-1 or 1). How do we compute this prediction?

1.1 Geometric Intuition for LaGrange Multipliers

Solve the following constrained optimization problem:

$$\min_{x,y} f(x, y) = x^2 + y^2$$

Subject to

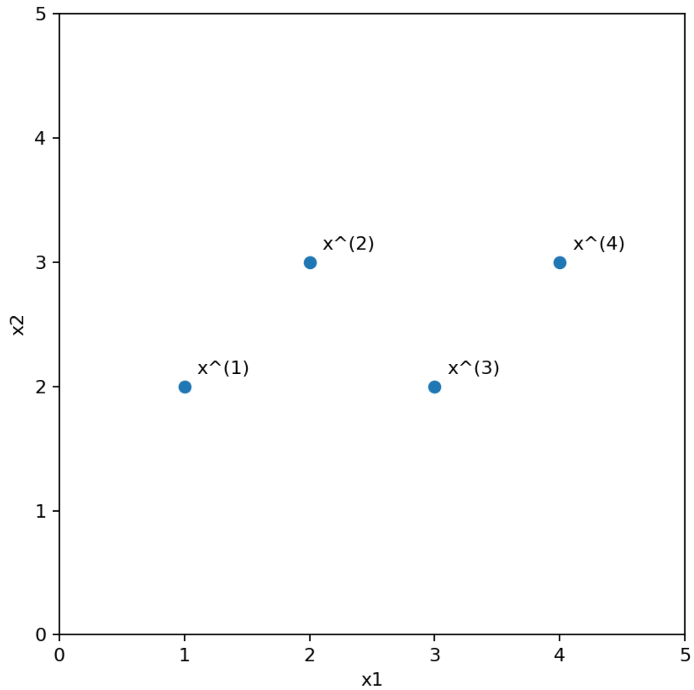
$$g(x, y) = 0$$

where

$$g(x, y) = x + y - 1$$

2 PCA: Basic Concepts

Consider dataset $\mathcal{D} = \{\mathbf{x}^{(1)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{x}^{(2)} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathbf{x}^{(3)} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{x}^{(4)} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}\}$. A visualization of the dataset is as below.



2.1 Centering Data

Centering is crucial for PCA. We must preprocess data so that all features have zero mean before applying PCA, i.e.

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{x}^{(i)} = \vec{0}$$

Compute the centered dataset:

$\mathbf{x}^{(1)} =$ __________ $\mathbf{x}^{(2)} =$ __________

$\mathbf{x}^{(3)} =$ __________ $\mathbf{x}^{(4)} =$ __________

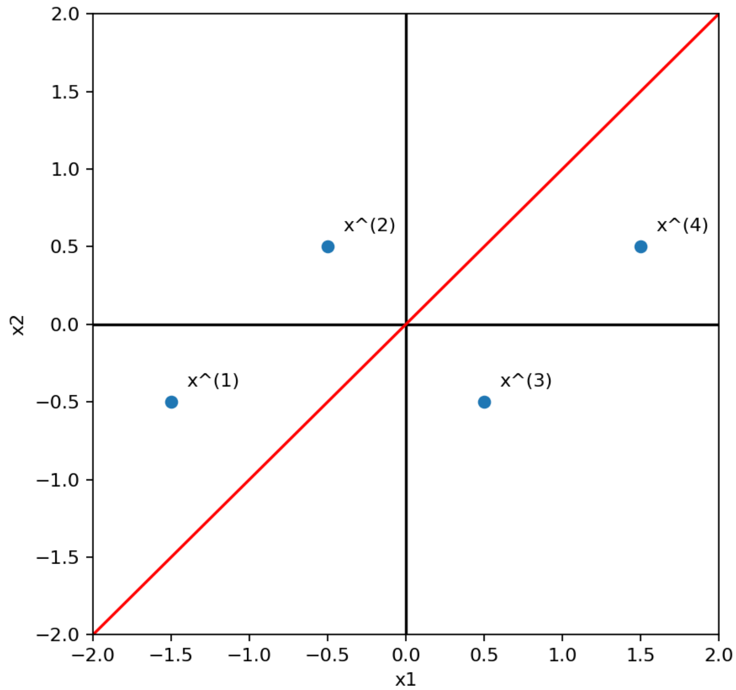
2.2 Unit vector

In order to easily compute the projected coordinates of data, we need to make the projected directions unit vectors. Suppose we want to project our data onto the vector $\mathbf{v} = \begin{bmatrix}1\\1\end{bmatrix}$. Normalize $\mathbf{v}$ to be a unit vector.

$\mathbf{v} =$ __________

2.3 Project Data

The centered data should now look like the following:



Suppose we want to project the centered data onto $\mathbf{v}$, where $\mathbf{v}$ goes through the origin.

Compute the magnitude of the projections, i.e. compute $z^{(i)} = \mathbf{v}^T\mathbf{x}^{(i)}, \forall 1 \leq i \leq N$.

$z^{(1)} =$ ____________ $z^{(2)} =$ ____________

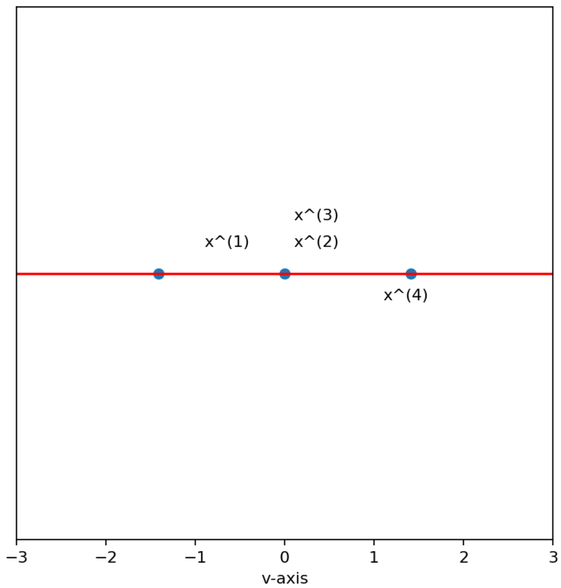
$z^{(3)} =$ ____________ $z^{(4)} =$ ____________

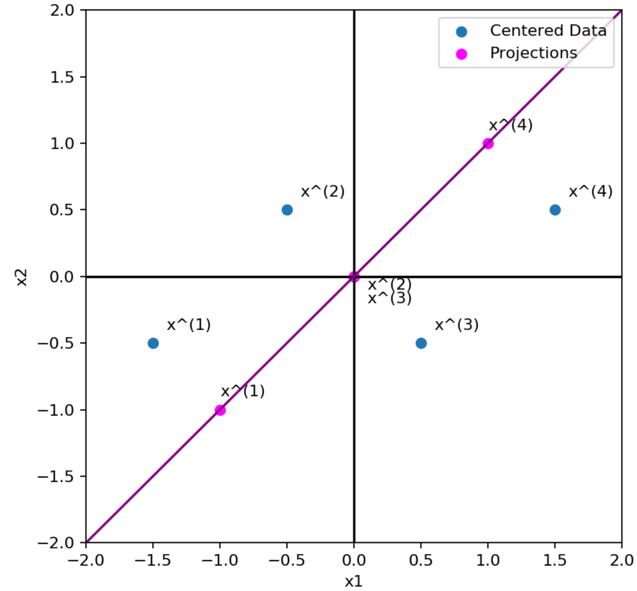
Let $\mathbf{x}^{(i)'}$ be the projected point of $\mathbf{x}^{(i)}$. Note that $\mathbf{x}^{(i)'} = \mathbf{v}^T\mathbf{x}^{(i)}\mathbf{v} = z^{(i)}\mathbf{v}$. Compute the projected coordinates:

$\mathbf{x}^{(1)'} =$ ____________ $\mathbf{x}^{(2)'} =$ ____________

$\mathbf{x}^{(3)'} =$ ____________ $\mathbf{x}^{(4)'} =$ ____________

Below is a visualization of the projections:





2.4 Reconstruction Error

One of the two goals of PCA is to find new directions to project our dataset onto such that it **minimizes the reconstruction error**, where the reconstruction error is defined as following:

$$\text{Reconstruction Error} = \sum_{i=1}^{N} \|\mathbf{x}^{(i)'} - \mathbf{x}^{(i)}\|_2^2$$

What is the reconstruction error in our case?

Reconstruction Error = ____________

2.5 Variance of Projected Data

Another goal is to find new directions to project our dataset onto such that it **maximizes the variance of the projections**, where the variance of projections is defined as following:

$$\begin{aligned}
\text{variance of projection} &= \sum_{i=1}^{N} (z^{(i)} - \hat{E}[z])^2 \\
&= \sum_{i=1}^{N} (\mathbf{v}^T\mathbf{x}^{(i)} - \frac{1}{N}\sum_{j=1}^{N} \{\mathbf{v}^T\mathbf{x}^{(j)})^2 \\
&= \sum_{i=1}^{N} (\mathbf{v}^T\mathbf{x}^{(i)} - \mathbf{v}^T(\frac{1}{N}\sum_{j=1}^{N} \mathbf{x}^{(j)}))^2 \\
&= \sum_{i=1}^{N} (\mathbf{v}^T\mathbf{x}^{(i)} - \mathbf{v}^T\vec{0})^2 \\
&= \sum_{i=1}^{N} (\mathbf{v}^T\mathbf{x}^{(i)})^2
\end{aligned}$$

What is the variance of the projections?

variance = ____________