

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

# Perceptron

## The simplest neural network

Dr. Mauricio Toledo-Acosta

Diplomado Ciencia de Datos con Python

# Table of Contents

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

1 Introducción

2 El método

3 El método dual

4 Kernel

# Introducción

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

## Perceptron

Modelo supervisado de clasificación binaria que busca encontrar una frontera de decisión que separe las clases de puntos. Esto lo hace usando una función umbral que depende de un vector de pesos, este vector lo va actualizando conforme recorre los puntos de entrenamiento.

# El modelo lineal de clasificación

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

- El Algoritmo Perceptrón (Rosenblatt, 1961) jugó un rol muy importante en la historia del Machine Learning. Inicialmente, simulado en una computadora IBM 704 en Cornell en 1957. Para principios de 1960, se había diseñado un hardware de propósito específico, que implementaba directamente y de forma paralela el algoritmo de aprendizaje.

# El modelo lineal de clasificación

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

- El Algoritmo Perceptrón (Rosenblatt, 1961) jugó un rol muy importante en la historia del Machine Learning. Inicialmente, simulado en una computadora IBM 704 en Cornell en 1957. Para principios de 1960, se había diseñado un hardware de propósito específico, que implementaba directamente y de forma paralela el algoritmo de aprendizaje.
- Fue criticado por Marvin Minsky, mostrando las deficiencias de las redes de perceptrón de una sola capa para resolver problemas no separables. Esto propició un vacío en la investigación de la computación neural que duró hasta mediados de los 80.

# Table of Contents

Perceptron

**Clasificación**

Introducción

El método

El método  
dual

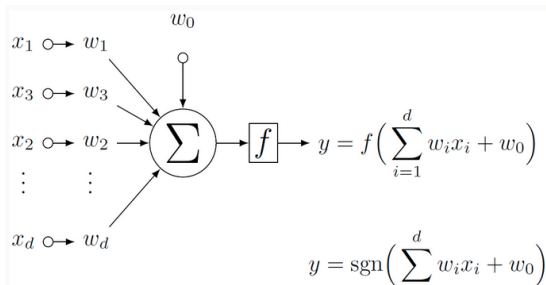
Kernel

1 Introducción

2 El método

3 El método dual

4 Kernel



# El algoritmo

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

Recorremos el conjunto de entrenamiento

$$(x_1, y_1), (x_2, y_2) \cdots, (x_N, y_N).$$



# El algoritmo

Perceptron

Clasificación

Introducción

El método

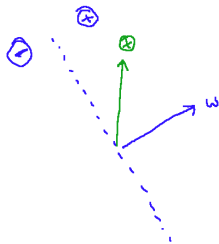
El método  
dual

Kernel

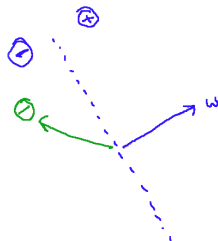
Recorremos el conjunto de entrenamiento

$$(x_1, y_1), (x_2, y_2) \cdots, (x_N, y_N).$$

- Si el punto está bien clasificado:



(a) Si  $y = +1$



(b) Si  $y = -1$

# El algoritmo

Perceptron

Clasificación

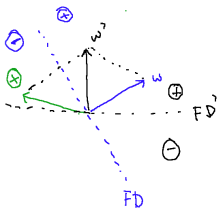
Introducción

El método

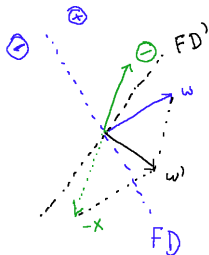
El método  
dual

Kernel

- Si el punto está mal clasificado:



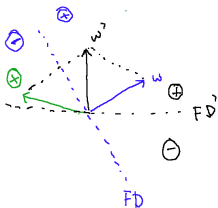
(a) Si  $y = +1$



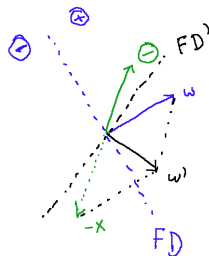
(b) Si  $y = -1$

## El algoritmo

- b** Si el punto está mal clasificado:



(a) Si  $y = +1$



(b) Si  $y = -1$

Actualizamos el vector  $w$  con  $w + yx$

# El algoritmo

Perceptron

Clasificación

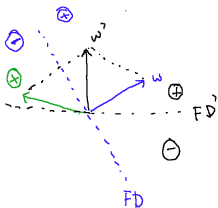
Introducción

El método

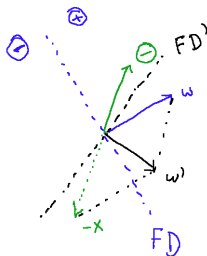
El método dual

Kernel

- Si el punto está mal clasificado:



(a) Si  $y = +1$



(b) Si  $y = -1$

Actualizamos el vector  $w$  con  $w + yx$

Repetimos para el conjunto de entrenamiento y realizamos varias épocas.

# El algoritmo

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

El algoritmo se resume así, para cada época:

**Entrada:** Datos etiquetados de entrenamiento  $\mathbf{X}$  en coordenadas homogéneas

**Salida:** Vector de pesos  $\mathbf{w}$  que define al clasificador

$\mathbf{w} = \mathbf{0}$  #Otras inicializaciones son posibles

$converge = \text{Falso}$

**mientras**  $converge == \text{Falso}$  :

$converge = \text{Verdadero}$

**para**  $i$  en  $|\mathbf{X}|$  :

**si**  $y_i \langle \mathbf{w}, \mathbf{x}_i \rangle \leq 0$  **entonces:** #xi mal clasificado

$\mathbf{w} = \mathbf{w} + y_i \eta \mathbf{x}_i$  # $0 < \eta \leq 1$  es la tasa de aprendizaje

$converge = \text{Falso}$

**fin**

**fin**

**fin**

# El algoritmo

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

El algoritmo se resume así, para cada época:

**Entrada:** Datos etiquetados de entrenamiento  $\mathbf{X}$  en coordenadas homogéneas

**Salida:** Vector de pesos  $\mathbf{w}$  que define al clasificador

$\mathbf{w} = \mathbf{0}$  #Otras inicializaciones son posibles

$converge = \text{Falso}$

**mientras**  $converge == \text{Falso}$  :

$converge = \text{Verdadero}$

**para**  $i$  en  $|\mathbf{X}|$  :

**si**  $y_i \langle \mathbf{w}, \mathbf{x}_i \rangle \leq 0$  **entonces:** # $x_i$  mal clasificado

$\mathbf{w} = \mathbf{w} + y_i \eta \mathbf{x}_i$  #  $0 < \eta \leq 1$  es la tasa de aprendizaje

$converge = \text{Falso}$

**fin**

**fin**

**fin**

El vector  $\mathbf{w}$  es una combinación lineal de los vectores  $\mathbf{x}$ .

# Table of Contents

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

- 1 Introducción
- 2 El método
- 3 El método dual
- 4 Kernel

# El método dual

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

Observar que, al final del entrenamiento:

$$w = \alpha_1 y_1 x_1 + \alpha_2 y_2 x_2 + \cdots \alpha_N y_N x_N,$$

donde los  $\alpha_i$  tienen que ver con el número de veces que el punto  $x_i$  fue mal clasificado.

Entonces,  $w = \sum_{i=1}^N \alpha_i y_i x_i$  y por lo tanto

$$y_j = \text{sign} \left( \sum_{i=1}^N \alpha_i y_i \langle x_j, x_i \rangle \right)$$



# Algoritmo del perceptron dual

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

$\alpha := (\alpha_1, \alpha_2, \dots, \alpha_N) = 0$

*converge* = **Falso**

**mientras** *converge* == **Falso** :

*converge* = **Verdadero**

**para**  $i$  **en**  $|\mathbf{X}|$  :

toma  $(\mathbf{x}_i, y_i)$  de los datos

**si**  $y_i \sum_{j=1}^{|\mathbf{X}|} \alpha_j y_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \leq 0$  **entonces**: #xi mal clasificado

$\alpha_i = \alpha_i + 1$

*converge* = **Falso**

**fin**

# Table of Contents

Perceptron

**Clasificación**

Introducción

El método

El método  
dual

Kernel

1 Introducción

2 El método

3 El método dual

4 Kernel

# El truco del Kernel

Perceptron

**Clasificación**

Introducción

El método

El método  
dual

Kernel



# El truco del Kernel

Perceptron

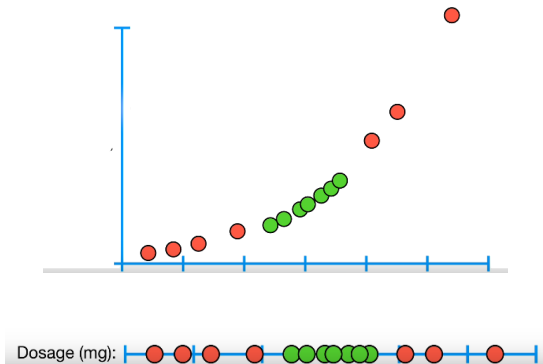
Clasificación

Introducción

El método

El método  
dual

Kernel



# Ejemplo $z = x^2 + y^2$

Perceptron

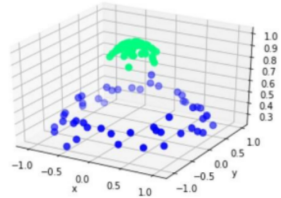
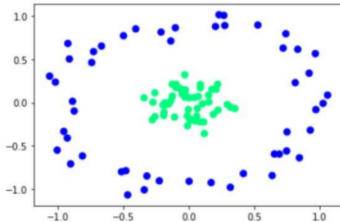
Clasificación

Introducción

El método

El método  
dual

Kernel



# Tipos de Kernel

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

- Aunque nuevos kernels aparecen en la literatura, los siguientes cuatro son básicos y ampliamente utilizados:

<b>Lineal:</b>	$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$
<b>Polinomial:</b>	$\kappa(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle + r)^p, r \geq 0$
<b>Gaussiano (<i>Radial Basis Function</i> – RBF):</b>	$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{-\ \mathbf{x}_i - \mathbf{x}_j\ ^2}{2\sigma^2}\right) = \exp\left(-\gamma\ \mathbf{x}_i - \mathbf{x}_j\ ^2\right)$
<b>Sigmoide:</b>	$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma\langle \mathbf{x}_i, \mathbf{x}_j \rangle + r)$

donde  $r, p, \gamma$  son parámetros de los modelos.

# Algoritmo Perceptron con Kernel

Perceptron

Clasificación

Introducción

El método

El método  
dual

Kernel

$\alpha := (\alpha_1, \alpha_2, \dots, \alpha_N) = 0$  ;  $\kappa(\mathbf{x}_i, \mathbf{x}_j)$ : entrada

*converge* = **Falso**

**mientras** *converge* == **Falso** :

*converge* = **Verdadero**

**para**  $i$  en  $|\mathbf{X}|$  :

toma  $(\mathbf{x}_i, y_i)$  de los datos

**si**  $y_i \sum_{j=1}^{|\mathbf{X}|} \alpha_j y_j \kappa(\mathbf{x}_i, \mathbf{x}_j) \leq 0$  **entonces**: #xi mal clasificado

$\alpha_i = \alpha_i + 1$

*converge* = **Falso**