> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Linear Classification Models

Fabio A. González Ph.D.

Depto. de Ing. de Sistemas e Industrial Universidad Nacional de Colombia, Bogotá

April 9, 2018

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

1 Introduction

2 The perceptron

3 Logistic regression



Content

▲ロト ▲周ト ▲ヨト ▲ヨト ヨー のくで

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

1 Introduction

2 The perceptron

3 Logistic regression

4 Logistic regression optimization

Outline

・ロット 4回ット 4回ット 4回ット 4日ッ

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regressior

Logistic regression optimization

Classification problems

▲□▶ ▲圖▶ ▲ 臣▶ ▲ 臣▶ ― 臣 … のへぐ

•
$$predict(x) = \begin{cases} C_1, & y(x) \ge \text{threshold} \\ C_2, & y(x) < \text{threshold} \end{cases}$$
,
with threshold = 0 or threshold = 0.5.

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Classification problems

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

•
$$predict(x) = \begin{cases} C_1, & y(x) \ge \text{threshold} \\ C_2, & y(x) < \text{threshold} \end{cases}$$
,
with threshold = 0 or threshold = 0.5.

• Three ways to address the classification problem:

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Classification problems

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

•
$$predict(x) = \begin{cases} C_1, & y(x) \ge \text{threshold} \\ C_2, & y(x) < \text{threshold} \end{cases}$$
,
with threshold = 0 or threshold = 0.5.

- Three ways to address the classification problem:
 - 1 Directly model the discrimination function: e.g $y(x) = w^T x + w_0$

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Classification problems

•
$$predict(x) = \begin{cases} C_1, & y(x) \ge \text{threshold} \\ C_2, & y(x) < \text{threshold} \end{cases}$$
,
with threshold = 0 or threshold = 0.5.

- Three ways to address the classification problem:
 - 1 Directly model the discrimination function: e.g $y(x) = w^T x + w_0$
 - 2 Generative model:

$$y(x) = P(C_k|x) = \frac{P(x|C_k)P(C_k)}{P(x)}$$

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Classification problems

•
$$predict(x) = \begin{cases} C_1, & y(x) \ge \text{threshold} \\ C_2, & y(x) < \text{threshold} \end{cases}$$
,
with threshold = 0 or threshold = 0.5.

- Three ways to address the classification problem:
 - 1 Directly model the discrimination function: e.g $y(x) = w^T x + w_0$
 - 2 Generative model:

$$y(x) = P(C_k|x) = \frac{P(x|C_k)P(C_k)}{P(x)}$$

3 Discriminative model:

$$y(x) = P(C_k|x) = f(x),$$

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

with f an arbitrary function.

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

Introduction

Linear Classification

> Models Fabio A. González Ph.D.

The perceptror

Logistic regression

Logistic regression optimization

$$y(x) = f(w^T x + w_0)$$

• $f(\cdot)$: activation function, may be non-linear

▲□▶▲□▶▲≡▶▲≡▶ ≡ めぬぐ

Ph.D.

Linear Classification

> Models Fabio A. González

The perceptror

Logistic regressior

$$y(x) = f(w^T x + w_0)$$

- $f(\cdot)$: activation function, may be non-linear
- Even if $f(\cdot)$ is non-linear, the decision boundary is linear

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

Ph.D.

Linear Classification

> Models Fabio A. González

The perceptron

Logistic regression

$$y(x) = f(w^T x + w_0)$$

- $f(\cdot)$: activation function, may be non-linear
- Even if $f(\cdot)$ is non-linear, the decision boundary is linear
- Also called generalized linear models

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Ph.D.

Linear Classification

> Models Fabio A. González

The perceptron

Logistic regression

$$y(x) = f(w^T x + w_0)$$

- $f(\cdot)$: activation function, may be non-linear
- Even if $f(\cdot)$ is non-linear, the decision boundary is linear
- Also called generalized linear models
- Applicable if instead of x we use a vector of basis functions $\phi(x)$, corresponding to features in a feature space

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regressior

Logistic regression optimization

Using regression for classification

• We can use a regression model, such as least squares to fit a linear classification model with a linear activation function

$$\min_{w,w_o} \sum_{i=1}^{\ell} (t_i - w^T x_i + w_0)^2,$$

where $t_i \in \{-1, 1\}$ is the label of the *i*-th training sample,

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Using regression for classification

• We can use a regression model, such as least squares to fit a linear classification model with a linear activation function

$$\min_{w,w_o} \sum_{i=1}^{\ell} (t_i - w^T x_i + w_0)^2,$$

where $t_i \in \{-1, 1\}$ is the label of the *i*-th training sample,

• but this strategy does not work well:



▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

1 Introduction

2 The perceptron

3 Logistic regression

4 Logistic regression optimization

Outline

・ロト・日本・山田・山田・山田・

人口 医水黄 医水黄 医水黄素 化甘油



• Designed by Frank Rossemblat in 1957

Linear Classification Models

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @



- Designed by Frank Rossemblat in 1957
- A hardware implementation of the learning algorithm

Linear Classification Models

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ ▲ 三 ● ● ●



- Designed by Frank Rossemblat in 1957
- A hardware implementation of the learning algorithm
- The precursor of neural networks

Linear Classification Models

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression



- Designed by Frank Rossemblat in 1957
- A hardware implementation of the learning algorithm
- The precursor of neural networks
- Criticized by Marvin Minsky, producing a decline in research funding

Linear Classification Models

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regressior

Logistic regression optimization

Perceptron learning

▲□▶ ▲圖▶ ▲ 臣▶ ▲ 臣▶ ― 臣 … のへぐ

• Activation function:

$$f(a) = \begin{cases} +1, & a \ge 0\\ -1, & a < 0 \end{cases}$$

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Perceptron learning

▲□▶ ▲圖▶ ▲ 臣▶ ▲ 臣▶ ― 臣 … のへぐ

• Activation function:

$$f(a) = \begin{cases} +1, & a \ge 0\\ -1, & a < 0 \end{cases}$$

• Loss function:

$$E_p(w, w_0) = -\sum_{n=1}^{\ell} f(w^T x_n + w_0) t_n,$$

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Perceptron learning

• Activation function:

$$f(a) = \begin{cases} +1, & a \ge 0\\ -1, & a < 0 \end{cases}$$

• Loss function:

$$E_p(w, w_0) = -\sum_{n=1}^{\ell} f(w^T x_n + w_0) t_n,$$

• Learning rule:

$$w^{(n)} = w^{(n-1)} + \eta (f_n - t_n) x_n$$

▲□▶▲□▶▲≡▶▲≡▶ ≡ めぬぐ

where $f_n = f(w^T x_n + w_0)$

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Perceptron convergence

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

• If the training points are linearly separable, the perceptron algorithms converges (*perceptron convergence theorem*)

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Perceptron convergence

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

- If the training points are linearly separable, the perceptron algorithms converges (*perceptron convergence theorem*)
- It could converge to different solutions depending on the order of presentation of training sample

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Perceptron convergence

э

- If the training points are linearly separable, the perceptron algorithms converges (*perceptron convergence theorem*)
- It could converge to different solutions depending on the order of presentation of training sample



Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Perceptron problems

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 - のへで

• Non-probabilistic outputs

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Perceptron problems

▲ロト ▲周ト ▲ヨト ▲ヨト ヨー のくで

- Non-probabilistic outputs
- Non-convex optimization problem

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Perceptron problems

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

- Non-probabilistic outputs
- Non-convex optimization problem
- No convergence guarantee if samples are not linearly separable

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Perceptron problems

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

- Non-probabilistic outputs
- Non-convex optimization problem
- No convergence guarantee if samples are not linearly separable
- Its power may be increased by stacking several perceptron layers (multilayer perceptrons) and using smooth activation functions

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

1 Introduction

2 The perceptron



4 Logistic regression optimization

Outline

・ロト・西ト・西ト・日・ 日・ シック

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Parametric discrimination

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

• These three conditions are equivalent:

- $P(C_1|x) \ge 0.5$
- $\frac{P(C_1|x)}{1-P(C_1|x)} \ge 1$
- $\operatorname{logit}(P(C_1|x)) = \log \frac{P(C_1|x)}{1 P(C_1|x)} \ge 0$

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Parametric discrimination

• These three conditions are equivalent:

- $P(C_1|x) \ge 0.5$ • $\frac{P(C_1|x)}{1-P(C_1|x)} \ge 1$
- $\frac{1-P(C_1|x)}{1-P(C_1|x)} \ge 1$ • $\operatorname{logit}(P(C_1|x)) = \log \frac{P(C_1|x)}{1-P(C_1|x)} \ge 0$
- If we assume that $P(x|C_1)$ and $P(x|C_2)$ are normally distributed sharing the same covariance matrix:

$$logit(P(C_1|x)) = log \frac{P(C_1|x)}{P(C_2|x)} = w^T x + w_0,$$

where

$$w = \Sigma^{-1}(\mu_1 + \mu_2)$$

$$w_0 = -\frac{1}{2}(\mu_1 + \mu_2)^T \Sigma^{-1}(\mu_1 + \mu_2) + \log \frac{P(C_1)}{P(C_2)}$$

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Logistic function

▲□▶ ▲圖▶ ▲ 臣▶ ▲ 臣▶ ― 臣 … のへぐ

• The logit function:

$$logit(P(C_1|x)) = log \frac{P(C_1|x)}{1 - P(C_1|x)} = w^T x + w_0$$

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Logistic function

▲□▶ ▲圖▶ ▲ 臣▶ ▲ 臣▶ ― 臣 … のへぐ

• The logit function:

$$logit(P(C_1|x)) = log \frac{P(C_1|x)}{1 - P(C_1|x)} = w^T x + w_0$$

$$P(C_1|x) = \sigma(w^T x + w_0) = \frac{1}{1 + e^{-(w^T x + w_0)}}$$

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Logistic function

• The logit function:

$$logit(P(C_1|x)) = log \frac{P(C_1|x)}{1 - P(C_1|x)} = w^T x + w_0$$

• The inverse-logit:

$$P(C_1|x) = \sigma(w^T x + w_0) = \frac{1}{1 + e^{-(w^T x + w_0)}}$$

• σ is called the logistic or sigmoid function.



> Fabio A. González Ph.D.

Introduction

The perceptror

Logistic regression

Logistic regression optimization

Logistic regression

◆□ > ◆□ > ◆ 三 > ◆ 三 > ● ○ ○ ○ ○

$$y(x) = P(C_1|x) = \sigma(w^T x)$$

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Logistic regression

▲□▶▲□▶▲≡▶▲≡▶ ≡ めぬぐ

$$y(x) = P(C_1|x) = \sigma(w^T x)$$

• Find w using maximum likelihood estimation:

$$p(\mathbf{t}|w) = \prod_{n=1}^{\ell} y_n^{t_n} (1-y_n)^{1-t_n},$$

where $\mathbf{t} = \{t_1, \ldots, t_\ell\}$ and $y_n = y(x_n)$.

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Logistic regression

$$y(x) = P(C_1|x) = \sigma(w^T x)$$

• Find w using maximum likelihood estimation:

$$p(\mathbf{t}|w) = \prod_{n=1}^{\ell} y_n^{t_n} (1 - y_n)^{1 - t_n},$$

where
$$\mathbf{t} = \{t_1, \dots, t_\ell\}$$
 and $y_n = y(x_n)$.

• Cross-entropy error:

$$E(w) = -\ln p(\mathbf{t}|w) = -\sum_{n=1}^{\ell} \left[t_n \ln y_n + (1 - t_n) \ln(1 - y_n) \right]$$

◆□▶ ◆圖▶ ◆臣▶ ◆臣▶ 三臣 - 釣��



Fabio A. González Ph.D.

Introduction

The perceptro

Logistic regression

Logistic regression optimization

Multiclass logistic regression

 $y_k(x) = P(C_k|x) = \frac{e^{w_k^T x}}{\sum_j e^{w_j^T x}}$

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

> Fabio A. González Ph.D.

Introduction

The perceptror

Logistic regression

Logistic regression optimization

Multiclass logistic regression

$$y_k(x) = P(C_k|x) = \frac{e^{w_k^T x}}{\sum_j e^{w_j^T x}}$$

• Likelihood:

$$p(\mathbf{T}|w_1\ldots w_K) = \prod_{n=1}^{\ell} \prod_{k=1}^{K} y_{nk}^{t_{nk}},$$

where $y_{nk} = y_k(x_n)$ and $\mathbf{T} \in \mathbb{R}^{\ell \times K}$ is a matrix of target variables with elements t_{nk} .

▲□▶ ▲□▶ ▲□▶ ▲□▶ ■ ●の00

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Multiclass logistic regression

$$y_k(x) = P(C_k|x) = \frac{e^{w_k^T x}}{\sum_j e^{w_j^T x}}$$

• Likelihood:

$$p(\mathbf{T}|w_1\ldots w_K) = \prod_{n=1}^{\ell} \prod_{k=1}^{K} y_{nk}^{t_{nk}},$$

where $y_{nk} = y_k(x_n)$ and $\mathbf{T} \in \mathbb{R}^{\ell \times K}$ is a matrix of target variables with elements t_{nk} .

• Multiclass cross-entropy error:

$$E(w_1, \dots, w_K) = -\ln p(\mathbf{T}|w_1 \dots w_K) = -\sum_{n=1}^{\ell} \sum_{k=1}^{K} t_{nk} \ln y_{nk}$$

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

1 Introduction

The perceptron

3 Logistic regression



Outline

・ロト・日本・日本・日本・日本・日本

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Optimization problem

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

$$\min_{w} E(w) = \min_{w} - \sum_{n=1}^{\ell} \left[t_n \ln y_n + (1 - t_n) \ln(1 - y_n) \right]$$

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Optimization problem

$$\min_{w} E(w) = \min_{w} -\sum_{n=1}^{\ell} \left[t_n \ln y_n + (1 - t_n) \ln(1 - y_n) \right]$$

$$\nabla E(w) = \sum_{n=1}^{\ell} (y_n - t_n)\phi_n$$

◆□ ▶ ◆□ ▶ ◆ 臣 ▶ ◆ 臣 ▶ ○ 臣 ○ のへで

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Optimization problem

$$\min_{w} E(w) = \min_{w} -\sum_{n=1}^{\ell} \left[t_n \ln y_n + (1 - t_n) \ln(1 - y_n) \right]$$

$$\nabla E(w) = \sum_{n=1}^{\ell} (y_n - t_n)\phi_n$$

$$w^{(\tau+1)} = w^{(\tau)} - \eta \sum_{n=1}^{\ell} (y_n - t_n)\phi_n$$

0

◆□ ▶ ◆□ ▶ ◆ 臣 ▶ ◆ 臣 ▶ ○ 臣 ○ のへで

Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Newton-Raphson

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

$$w^{(\tau+1)} = w^{(\tau)} - \mathbf{H}^{-1} \nabla E(w)$$

Fabio A. González Ph.D.

Introduction

The perceptror

Logistic regressior

Logistic regression optimization

Newton-Raphson

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

$$w^{(\tau+1)} = w^{(\tau)} - \mathbf{H}^{-1} \nabla E(w)$$

$$\nabla E(w) = \Phi^T(\mathbf{y} - \mathbf{t})$$

Newton-Raphson

m

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

$$w^{(\tau+1)} = w^{(\tau)} - \mathbf{H}^{-1} \nabla E(w)$$
$$\nabla E(w) = \Phi^{T}(\mathbf{y} - \mathbf{t})$$

$$\mathbf{H} =
abla
abla E(w) = \Phi^T \mathbf{R} \Phi,$$

with \mathbf{R} a diagonal matrix with $R_{nn} = y_n(1-y_n).$

The perceptron

Linear Classification

Models Fabio A. González Ph.D.

Logistic regression

Newton-Raphson

▲ロ ▶ ▲周 ▶ ▲ 国 ▶ ▲ 国 ▶ ● の Q @

$$w^{(\tau+1)} = w^{(\tau)} - \mathbf{H}^{-1} \nabla E(w)$$
$$\nabla E(w) = \Phi^{T}(\mathbf{y} - \mathbf{t})$$

$$\mathbf{H} = \nabla \nabla E(w) = \Phi^T \mathbf{R} \Phi,$$

with **R** a diagonal matrix with $R_{nn} = y_n(1 - y_n)$.

Linear Classification

> Models Fabio A. González Ph.D.

Logistic regression

optimization

• The resulting algorithm is called *iterative reweighted least* squares.

Regularization

◆□▶ ◆□▶ ◆ 臣▶ ◆ 臣▶ ○ 臣 ○ の Q @



Linear Classification

Models Fabio A. González

The perceptro

Logistic regression

$$\min_{w} - C \sum_{n=1}^{\ell} \left[t_n \ln y_n + (1 - t_n) \ln(1 - y_n) \right] + \|w\|^2$$

Regularization

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

Fabio A. González Ph.D.

Linear Classification

Models

Introduction

The perceptro

Logistic regressior

Logistic regression optimization

$$\min_{w} - C \sum_{n=1}^{\ell} \left[t_n \ln y_n + (1 - t_n) \ln(1 - y_n) \right] + \|w\|^2$$

• Prevents overfitting.

Regularization

▲□▶▲□▶▲≡▶▲≡▶ ≡ めぬぐ

Ph.D.

Linear Classification

> Models Fabio A. González

The perceptro

Logistic regressior

$$\min_{w} - C \sum_{n=1}^{\ell} \left[t_n \ln y_n + (1 - t_n) \ln(1 - y_n) \right] + \|w\|^2$$

- Prevents overfitting.
- Equivalent to the inclusion of a prior and finding a MAP solution for *W*.

Fabio A. González Ph.D.

Introduction

The perceptror

Logistic regressior

Logistic regression optimization

Stochastic gradient descent

$$\min_{w} Q(w) = \min_{w} \sum_{i=1}^{n} Q_i(w)$$

◆□▶ ◆□▶ ◆ □▶ ◆ □▶ ○ □ ○ ○ ○ ○

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Stochastic gradient descent

$$\min_{w} Q(w) = \min_{w} \sum_{i=1}^{n} Q_i(w)$$

• Batch gradient descent:

$$w^{(\tau+1)} = w^{(\tau)} - \alpha \nabla Q(w) = w^{(\tau)} - \alpha \sum_{i=1}^{n} \nabla Q_i(w)$$

▲□▶ ▲□▶ ▲ 三▶ ▲ 三▶ 三三 - のへぐ

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization

Stochastic gradient descent

$$\min_{w} Q(w) = \min_{w} \sum_{i=1}^{n} Q_i(w)$$

• Batch gradient descent:

$$w^{(\tau+1)} = w^{(\tau)} - \alpha \nabla Q(w) = w^{(\tau)} - \alpha \sum_{i=1}^{n} \nabla Q_i(w)$$

• on-line gradient descent:

$$w^{(\tau+1)} = w^{(\tau)} - \alpha \nabla Q_i(w)$$

▲□▶ ▲□▶ ▲□▶ ▲□▶ □ のQ@

> Fabio A. González Ph.D.

Introduction

The perceptron

Logistic regression

Logistic regression optimization Alpaydin, E. 2010 Introduction to Machine Learning, 2Ed. The MIT Press. (Chap 10)

Russell, S and Norvig, P. 2010 Artificial Intelligence: a Modern Approach, 3rd Ed, Prentice-Hall (Sect 18.6)

▲ロト ▲周ト ▲ヨト ▲ヨト ヨー のくで