Lab 4 – Logistic Regression

We started from...

$$min_{w \in \mathbb{R}^D} \frac{1}{n} \sum_{i=1}^{n} L(y_i, f_w(x_i)) + \lambda \|w\|^2$$

... and picked a specific loss function ...

$$L(y_i, f_w(x_i)) = log(1 + e^{-y_i f_w(x_i)})$$

Logistic Regression

$$min_{w \in \mathbb{R}^D} \frac{1}{n} \sum_{i=1}^{n} log(1 + e^{-y_i f_w(x_i)}) + \lambda ||w||^2$$

Despite the name we are solving a classification problem

Solving LR with Gradient Descent

$$w_t = w_{t-1} - \gamma \left(\frac{1}{n} \sum_{i=1}^{n} x_i \frac{-y_i}{1 + e^{y_i x_i^T w_{t-1}}} + 2\lambda w_{t-1} \right)$$

How to compute it?

How to initialize it?

When do we stop?

Solving LR with Stochastic Gradient Descent

$$w_t = w_{t-1} - \gamma_t \left(\frac{1}{n} \sum_{i=1}^{n} x_i \frac{-y_i}{1 + e^{y_i x_i^T w_{t-1}}} + 2\lambda w_{t-1} \right)$$

The notion of confidence

$$\frac{1}{1 + e^{-f_w(x)}}$$

$$= \frac{1}{1 + e^{-w^T x}}$$

Your objectives today

- Implementing Logistic Regression with Gradient Descent

- More specifically you will implement

  - The function **train_logreg_gd (Xtr, Ytr, reg_par)** to estimate the **w** using gradient descent

  - The function **predict_logreg (w, Xts)** to evaluate the function on a set of points (obtaining the prediction)

- Reason on the confidence…

- Extend to Stochastic Gradient Descent

- IMPORTANT NOTE: we are not considering CV in this lab as we are focusing on the analysis… but we would need it for selecting the most appropriate regularization parameter!!!

UniGe

MaLGa