Année 2014/2015 UNIVERSITÉ TOULOUSE I CAPITOLE

Big Data

TP 1 **Convex optimisation**

Short exercice on duality

Exercice 1. We define

$$J(\theta) = \frac{\theta_1^2 + \theta_2^2 + \theta_3^2}{2}.$$

Constraints are

$$f_1(\theta) = \theta_1 + \theta_2 + 2\theta_3 = 1 \qquad \text{and} \qquad f_2(\theta) = \theta_1 + 4\theta_2 + 2\theta_3 = 3.$$

1. Describe this optimization problem within a matricial form.
2. Describe the Lagrangian function $L(\theta, \lambda)$.
3. Write the KKT condition
4. Compute the dual function $\mathcal{L}(\lambda)$ and obtain the value of the Lagrange multiplier.
5. Come back to the initial constrained problem.

Optimization with Matlab

Exercice 2.

We aim to numerically test the simplest method of gradient descent with a constant step size γ on the standard problem of the ridge regression. We work on the p dimensional space $\mathbb{R}^p$ and define

$$J(w) = \|y - w^t x\|_2^2 + \lambda \|w\|_2^2.$$

We consider $p > n$ (for instance $p = 100$) and assume that $X \sim \mathcal{U}_{[0,1]^p}$. The response Y is a real response value defined by

$$Y = \theta_0^t X + \epsilon.$$

where θ_0 is unknown and $\epsilon \sim \mathcal{N}(0, \sigma^2 I_p)$.

1. Generate an i.i.d. sample of size $n = 50$, with $p = 2n$ and θ_0 chosen as you want.
2. Compute theoretically the expression of the Ridge regression estimator built from the function J:
$$\hat{\theta}_n := \arg\min_{\theta \in \mathbb{R}^p} J(\theta)$$
3. Show that J is a convex function.

4. Compute the gradient of J. Is J β smooth? strongly convex?

5. What kind of step size could we use for the gradient descent algorithm?

6. What happens when $\lambda \longmapsto 0$?

7. Develop a Matlab code to solve the ridge regression estimation. Is the estimation convenient when $p >> n$?

8. Choose now a sparse vector θ_0. Does the estimated $\hat{\theta}$ is satisfactory and select a sparse vector?

Duality

Exercice 3.

We consider $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}_+^n$ and define the following problem

$$\min_{x \in \mathbb{R}^n} -\sum_{i=1}^{n} \log(\alpha_i + x_i) \qquad \text{s.t.} \qquad (x_1, \ldots, x_n) \in \mathcal{S}^{n-1}.$$

1. Write the Lagrangian and derive a dual problem. (We will use the notation λ for equality constraints, and μ for inequality constraints).

2. Write the KKT conditions.

3. Solve the dual problem with Matlab and come back to the initial one.