ECP 2ème année
Semestre 4, Electif 11, Année 2010
Mr DUMAS

NUMERICAL OPTIMIZATION AND APPLICATIONS

final examen, 3 hours

All Scilab programs must be sent at the adress `dumas@ann.jussieu.fr` *at the end of the exam with the subjet ECP2010 and with a file name of the type* `ECP2010-firstname-name.sci`.

The problem is in english but candidates can give a copy in french if they prefer.

EXERCICE 1

This exercice is intented to prove some convergence results of a given descent method with a line-search strategy. First, let recall some notations : let f a given function C^1 from $\mathbb{R}^n$ to $\mathbb{R}$, such that $\lim\limits_{||x|| \to +\infty} f(x) = +\infty$. Denote by g the gradient function of f defined from $\mathbb{R}^n$ to $\mathbb{R}^n$. Assume that g is Lipschitz continous on $\mathbb{R}^n$ with a Lipschitz constant equal to L.

A descent method is a sequence $(x_k)_{k \in \mathbb{N}}$ defined from an arbitrary initial point $x_0 \in \mathbb{R}^n$ with the formula :

$$x_{k+1} = x_k + t_k d_k$$

where d_k is a descent direction (that is such $(d_k, g(x_k)) < 0$) and t_k is the corresponding step size in this direction obtained from a linesearch strategy.

Here the step size is given by the Curry rule : $t_k = 0$ if $g(x_k) = 0$, else :

$$t_k = \inf\{t \geq 0, \quad h_k'(t) = 0,\ h_k(t) < h_k(0)\}$$

where for all $t \geq 0$, $h_k(t) = f(x_k + t d_k)$.

1. Prove that the Curry rule is valid, that is t_k exists for all $k \geq 0$. Draw a picture to show the value of this step size on a general example of a non-convex function.

2. The Z-condition is a stopping criterion for a given linesearch strategy and writes as :

$$f(x_{k+1}) \leq f(x_k) - C||g(x_k)||^2 \cos(\theta_k)$$

for a given $C > 0$ independant of k and where

$$\cos(\theta_k) = \frac{-(d_k, g(x_k))}{||g(x_k)||.||d_k||}$$

The objective here is to prove that the Curry step satisfies the Z-condition.

2 a) By using a Taylor-type equality, prove that for all $t \in [0, t_k]$

$$h(t_k) \leq h(t) \leq h(0) + t(d_k, g(x_k)) + t^2 L \frac{||d_k||^2}{2} \tag{1}$$

2b) Find the value of t, called t_{min} where the right hand side of the inequality (1) is minimal.

2c) Prove that

$$0 \leq (d_k, g(x_k)) + t_k L ||d_k||^2$$

and deduce that

$$f(x_{k+1}) \leq f(x_k) - \frac{(d_k, g(x_k))^2}{2L||d_k||^2}$$

2c) Conclude.

3. Prove that the Z-condition implies that

$$\sum_{k=1}^{+\infty} ||g(x_k)||^2 \cos^2(\theta_k) < +\infty$$

4. Assume that $d_k = g(x_k)$. What can you say about the corresponding descent method with the Curry step ?

EXERCICE 2

Particle Swarm Optimization is an evolutionnary algorithm which make evolve a population (called swarm) of elements in $\mathbb{R}^n$ (called particles) by adapting their velocity to the group in order to find the minimum of a given function $J : \mathbb{R}^n \to \mathbb{R}$. For $1 \leq i \leq N_{pop}$ and at a given generation called `gen`, denote :

(i) x_i : the current position of the i-th particle in the population,

(ii) v_i the velocity of the i-th particle,

(iii) p_i : the best position found by the i-th particle so far (that is where J is minimal),

(iv) p_g : the best position found by the swarm so far.

The swarm is randomly initialized. Then, to pass form generation `gen` to genetation `gen+1`, the following operation is made

$$v_i := w v_i + c_1 \phi_1 \otimes (p_i - x_i) + c_2 \phi_2 \otimes (p_g - x_i)$$

and $x_i := x_i + v_i$ where ϕ_1 and ϕ_2 are two random vectors uniformly chosen in $[0,1]^n$. The positive parameters w, c_1 and c_2 are respectively called inertia weight and acceleration coefficients. The symbol $\otimes$ denotes a point-wise vector multiplication. A maximal speed V_{max} for any particle of the swarm is also imposed.

1. Use Scilab to program the PSO algorithm described above.

2. Apply the PSO algorithm to the very simple example of 2 particles and $n = 1$, $J(x) = x^2$, and with $w = c_1 = c_2 = 0.7$. Draw the corresponding trajectory of the particles on the parabola.

3. Apply the PSO algorithm to the more complex example of the Rastrigin function on $[-5, 5]^2$ that is $J(x_1, x_2) = \sum_{i=1}^{2}(x_i^2 - \cos(2\pi x_i)) + 2$ and with 10 particles. What are your main observations on the convergence of the algorithm ?