OPTIMIZATION. HOMEWORK 9

OSCAR DALMAU

Comments:

- Please, follow the general guidelines provided at the beginning of the course.
- In this homework, we expect the student to draw conclusions about the experimental performance of algorithms. So, feel free to include all the analysis tools learned in statistics.

(1) Consider the following optimization problem.

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\mathbf{Q}\boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x}$$

where

$$\begin{aligned}\mathbf{Q} &= \mathbf{P}\mathbf{D}\mathbf{P}^T\\ \mathbf{P} &= \prod_{j=1}^{m}\mathbf{H}_j\\ \mathbf{H}_j &= \mathbf{I} - 2\frac{\boldsymbol{u}_j\boldsymbol{u}_j^T}{\boldsymbol{u}_j^T\boldsymbol{u}_j}\end{aligned}$$

$\mathbf{I}$ is the identity matrix, $\mathbf{H}_j$ are orthogonal Householder matrices, $\boldsymbol{u}_j$ are vectors with entries generated randomly in $(-1,1)$. $\mathbf{D}$ is a diagonal matrix whose i-th component is defined by

$$\log d_i = \left(\frac{i-1}{n-1}\right) ncond, \; i = 1, 2, \cdots, n$$

Note that the parameter $ncond$ is related with the condition number of $\mathbf{Q}$, for example, if $ncond = 0$ then $\mathbf{Q}$ is the identity matrix.

In order to define $\boldsymbol{b}$ we generate a solution $\boldsymbol{x}^*$ with components chosen randomly in the interval $(-1,1)$, then

$$\boldsymbol{b} = \mathbf{Q}\boldsymbol{x}^*$$

Implement the Conjugate gradient algorithm and apply this method to the previous optimization for $m = 3$, $n = 10000$ and $ncond \in \{2, 4, 6\}$. For each *ncond* generate 30 problems and compute the **average of iterations and time (in seconds)** to get the solution. Plot the **average of the norm of the gradient** (you can select the first and last iterations)

(2) Compare the Conjugate Gradient results with the **Steepest descent with exact step size** and the **Barzilai-Borwein gradient method**:

Algorithm 1 The Barzilai-Borwein gradient method

Require: $\boldsymbol{x}_0$, τ_g
$\lambda_0 = 0$, $\boldsymbol{y}_0 = \boldsymbol{x}_0$ $k = 0$
while $\|\boldsymbol{g}_k\| > \tau_g$ **do**
  Compute the Barzilai-Borwein step size α_k
  $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_k \boldsymbol{g}_k$
  $k = k + 1$
end while

In order to compute the Barzilai-Borwein step size α_k we can use one of the following formula for $k > 0$

$$\alpha_k = \frac{\boldsymbol{s}_{k-1}^T \boldsymbol{y}_{k-1}}{\boldsymbol{y}_{k-1}^T \boldsymbol{y}_{k-1}}$$

$$\alpha_k = \frac{\boldsymbol{s}_{k-1}^T \boldsymbol{s}_{k-1}}{\boldsymbol{s}_{k-1}^T \boldsymbol{y}_{k-1}}$$

where $\boldsymbol{s}_{k-1} = \boldsymbol{x}_k - \boldsymbol{x}_{k-1}$ and $\boldsymbol{y}_{k-1} = \boldsymbol{g}_k - \boldsymbol{g}_{k-1}$. If $k = 0$ use the exact step size.