C&O367: Nonlinear Optimization
(Winter 2013)
Assignment 1
H. Wolkowicz

Posted Sat. 2013-Jan-12

Due: Tues., 2013-Jan-22, 10:00AM (before class),

1 Local/Global Minimizers/Maximizers

1. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = 1000(x_2 - x_1^2)^2 + (1 - x_1)^2. \tag{1}$$

(a) Show that the point $x^* = \begin{pmatrix} 1 & 1 \end{pmatrix}^T$ is the only local minimum for f on $\mathbb{R}^2$ and, in fact is a strict local minimum.

Solution: We set the gradient to 0:

$$\nabla f(x) = \begin{pmatrix} -4000x_1(x_2 - x_1^2) - 2(1 - x_1) \\ 2000(x_2 - x_1^2) \end{pmatrix} = 0.$$

Then the second row implies that $x_2 - x_1^2 = 0$. The first row now implies $1 - x_1 = 0$ Both together now imply that x^* is the only critical (stationary) point. To check whether it is a strict local minimum we find the Hessian

$$\nabla^2 f(x) = \begin{bmatrix} -4000x_2 + 12000x_1^2 + 2 & -4000x_1 \\ -4000x_1 & 2000 \end{bmatrix}$$

and

$$\nabla^2 f(x^*) = \begin{bmatrix} 8002 & -4000 \\ -4000 & 2000 \end{bmatrix}$$

is positive definite by the leading principal minors test, i.e. $8002 > 0, ((8002)(2000) - (4000)(4000)) > 0$. Therefore x^* is indeed a strict local minimum.

(b) What can you say about x^* and the test for a global minimum? Is x^* a (strict) global minimum?

Solution: Note that the test for a global minimum fails since $\nabla^2 f(x)$ has a negative diagonal element whenever $-4000x_2 + 12000x_1^2 + 2 < 0$. However, $f(x^*) = 0$ and $f(x) \geq 0, \forall x$, clearly means that x^* is a global minimum. And it is clearly the only point where $f(x) = 0$ and so is a strict global minimum.

(c) Use the MATLAB *surf* and/or *mesh* command to provide plots of f for values $0 \le x_1 \le 2, 0 \le x_2 \le 2$ in steps of $.1$. (Hand in the plot(s), and in colour if possible. Include a printout of the commands you used.)

Solution:

```
[X,Y] = meshgrid(0:.1:2);
f= 1000*(Y-X.^2).^2  + (1-X).^2;
figure(1)
surf(X,Y,f)
figure(2)
mesh(X,Y,f)
```

(d) Documentation http://www.mathworks.com/help/optim/index.html is for the MATLAB toolbox.
Use the MATLAB function *fminunc* to minimize the function in equation (1) above. Use options and optimset so that your program first runs *without* using gradients and Hessians and then runs *with* using gradients and Hessians during the iterations. Hand in your program and the appropriate output parameters. Use initial starting points $(0\ 0)^T$ and $(-10\ \ -10)^T$.

Solution:

```
%options = optimset('Hessian','off');
options = optimset('Hessian','on');
options = optimset('GradObj','on');
x0=zeros(2,1);
[x,fval,exitflag,output,grad] = fminunc(@myfun,x0,options);
```

with function myfun:

```
function    [f,g,H] = myfun(x)
n=length(x);
f= 1000*(x(2)-x(1)^2)^2+(1-x(1))^2;
g=[ -4000*x(1)*(x(2)-x(1)^2)    - 2*(1-x(1))
     2000*(x(2)-x(1)^2)];
H= [ -4000*x(2)+12000*x(1)^2+2 -4000*x(1)
    -4000*x(1) 2000];
```

2. Find the local and global minimizers and maximizers of the following functions. (Carefully state your reasoning.)

(a) $f(x) = \frac{1}{5}x^5 - \frac{13}{4} * x^4 + \frac{59}{3}x^3 - \frac{107}{2}x^2 + 60x - 3$.

Solution: The derivative of f can be factored to get $f'(x) = (x-1) * (x-3) * (x-4) * (x-5)$ (One can use the symbolic toolbox in MATLAB.) The evaluation of the second derivative at the points $1, 3, 4, 5$ yield: $-24, 4, -3, 8$, respectively. Therefore the points $x = 1, 3, 4, 5$ are respectively local max, min, max, min. Since the degree of f is 5 is clear that it is unbounded above and unbounded below. Therefore, the local min/max are not global.

(b) $f(x) = x^2 e^{-x^2}$

Solution: We have $f'(x) = 2x(1-x^2)e^{-x^2}$. Hence, the critical points are $x = 0, -1, 1$. The evaluation of the second derivative at the points $0, -1, 1$ yield: $2, -4e^{-1}, -4e^{-1}$, respectively. Therefore the points $x = 0, -1, 1$ are respectively local min, max, max. Also not that $f(x) \geq 0$ for all points, so $x = 0$ is a global min. For global max, we have $f'(x) > 0$ for $x \in (-\infty, -1)$ and $x \in (0, 1)$, and $f'(x) < 0$ for $x \in (-1, 0)$ and $x \in (1, \infty)$. We also have $f(1) = f(-1)$, so $x = -1$ and $x = 1$ are global max.

2 Symmetric Matrices and Quadratic Forms

1. Classify the matrices according to whether they are positive or negative definite or semidefinite or indefinite.

(a) $\begin{bmatrix} 8 & 5 & 5 \\ 5 & 4 & 5 \\ 5 & 5 & 2 \end{bmatrix}$

Solution: Indefinite after checking eigenvalues or determinant. Also $\{1, 3\}$ minor is negative $16 - 25 < 0$.

(b) $\begin{bmatrix} 10 & 6 & 7 \\ 6 & 8 & 3 \\ 7 & 3 & 8 \end{bmatrix}$

Solution: This matirx is positive definite. You can prove that by computing the determinant of the leading principle minors. We have:

$$\Delta_1 = 10, \quad \Delta_2 = 44, \quad \Delta_3 = 122.$$

2. (a) Write the quadratic form associated with the matrix $\begin{bmatrix} 10 & 6 & 7 & 3 \\ 6 & 8 & 3 & 3 \\ 7 & 3 & 8 & -4 \\ 3 & 3 & -4 & -8 \end{bmatrix}$

Solution: $10x_1^2 + 8x_2^2 + 8x_3^2 - 8x_4^2 + 12x_1x_2 + 14x_1x_3 + 6x_1x_4 + 6x_2x_3 + 6x_2x_4 - 8x_3x_4$

(b) Write the following quadratic form in the form $x^T A x$ for an appropriate matrix A: $\quad x_1x_4 - 3x_2x_3 + 7x_3x_1 + 9x_3^2 - 3.2x_4x_3 + 3x_2x_4$.

Solution: $A = \begin{bmatrix} 0 & 0 & 3.5 & .5 \\ 0 & 0 & -1.5 & 1.5 \\ 3.5 & -1.5 & 9 & -1.6 \\ .5 & 1.5 & -1.6 & 0 \end{bmatrix}$

3 Critical Points

Show that although the origin is a critical point of the function $f(x) = x_1^5 - x_1 x_2^6$, it is neither a local maximizer nor a local minimizer of $f(x_1, x_2)$.

Solution: The gradient is $\nabla f(x) = \begin{pmatrix} 5x_1^4 - x_2^6 \\ -6x_1 x_2^5 \end{pmatrix}$, which is 0 at the origin. We would like to show that it is a saddle point, i.e., if the Hessian has a positive and a negative eigenvalue, then the corresponding eigenvectors will be directions of increase and decrease, respectively. More precisely, if $Hv = \lambda v$ with $\lambda > 0$, then $f(0 + \alpha v) = f(0) + \alpha^2 \frac{1}{2} v^T H v + o(\alpha^2)$ shows that v is a direction of increase. (A similar argument holds for the negative eigenvalue.)

Now the Hessian is

$$H = \begin{bmatrix} 20x_1^3 & -6x_2^5 \\ -6x_2^5 & -30x_1 x_2^4 \end{bmatrix}$$

But this is 0 at the origin and so this does not obviously help.

However, along the curve $x_1 = x_2^2$, we see that $f(x) < 0$ for x_2 small, while along the line $x_1 = x_2$, with $x_1 > 0$, we see that $f(x) > 0$ for x_2 small. Therefore, we do not have a local max or min.

4 Coercive Functions

Show carefully if the following functions are coercive or not.

1. $f(x, y, z) = \log(x^2 y^2 z^2) - x - y - z$

   Solution: This function is not coercive. To prove that, let's define $a = (r, \frac{1}{r}, 1)$ for $r > 0$. Clearly we have $\lim_{r\to\infty} \|a\| = +\infty$. However,

   $$f(a) = \log(\frac{r^2}{r^2}) - r - \frac{1}{r} - 1 = -r - \frac{1}{r} - 1,$$

   where goes to $-\infty$ when $r \to +\infty$. Hence, $f(x, y, z)$ is not coercive.

2. $f(x, y, z) = x^2 + y^2 + z^2 - \sin(xyz)$

   Solution: This function is coercive. First of all note that $|\sin(xyz)| \leq 1$ for all values of (x, y, z). When $\|(x, y, z)\| \to +\infty$, at least one of $|x|$, $|y|$, and $|z|$ goes to $+\infty$. Hence we clearly have

   $$\lim_{\|(x,y,z)\|\to+\infty} f(x, y, z) = +\infty$$