IC-32 # Optimal Control **Winter 2006**

Benoît Chachuat
ME C2 401, Ph: 33844, benoit.chachuat@epfl.ch

Problem Set #3 (With Corrections)

1. Consider the following NLP problem:

$$\min_{\mathbf{x}\in\mathbb{R}^3} f(\mathbf{x}) := x_1^2 + x_1x_2 + 2x_2^2 - 6x_1 - 2x_2 - 12x_3 \tag{1}$$

$$\begin{aligned} \text{s.t.} \quad & g_1(\mathbf{x}) := 2x_1^2 + x_2^2 \leq 15 \\ & g_2(\mathbf{x}) := x_1 - 2x_2 - x_3 \geq -3 \\ & x_1, x_2, x_3 \geq 0. \end{aligned} \tag{2}$$

(a) Find an optimal solution $\mathbf{x}^\star$ to this problem using the function `fmincon` in MATLAB®'s Optimization Toolbox:

- Make sure that the medium scale SQP algorithm, with Quasi-Newton update and line-search, is the selected solver in `fmincon`;
- Set the solution point tolerance, function tolerance and constraint tolerance in `fmincon` to 10^{-7};
- Specify the initial guess as $\mathbf{x}^0 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^\mathsf{T}$;
- Make sure that the inequality constraint g_2 is treated as a *linear* constraint by `fmincon`;
- Solve the NLP by using a finite-difference approximation of the gradients of the objective function and constraints first. Then, resolve the problem by providing explicit expressions for the gradients of the objective function and constraints (ask `fmincon` to check for the gradients before running the optimization in this latter case);
- Make sure that the solver terminated successfully in each case.

Solution. A possible implementation (with analytic expressions for the gradients) is as follows:

exm1.m

```
clear all
options = optimset('Display', 'iter', 'GradObj', 'on', ...
                   'GradConstr', 'on', 'DerivativeCheck', 'on', ...
                   'LargeScale', 'off', 'HessUpdate', 'bfgs', ...
                   'Diagnostics', 'on',  'TolX', 1e-7, ...
                   'TolFun', 1e-7, 'TolCon', 1e-7,
                   'MaxFunEval', 100, 'MaxIter', 100 )
x0 = [ 1; 1; 1 ];
xL = [ 0; 0; 0 ];
A  = [ -1 2 1 ];
b  = [ 3 ];

%Solve NLP Problem
[xopt, fopt, iout ] = fmincon( @exm1_fun, x0, A, b, [], [], xL, [], ...
                               @exm1_ctr, options );
```

exm1_fun.m

```
%%%%%%%%%%%%%% FUNCTION TO BE MINIMIZED %%%%%%%%%%%%%%
% Objective: f(x) := x1^2+x1*x2+2*x2^2-6*x1-2*x2-12*x3
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [f,df] = exm1_fun(x)
  f = x(1)^2+x(1)*x(2)+2*x(2)^2-6*x(1)-2*x(2)-12*x(3); % function
  if nargout > 1
    df = [ 2*x(1)+x(2)-6 % gradient
           x(1)+4*x(2)-2
           -12 ];
  end
end
```

exm1_ctr.m

```
%%%%%%%%%%%%%%%%%%%% CONSTRAINTS %%%%%%%%%%%%%%%%%%%%
% nonlin. ineq. constraint: g1(x) := 2*x1^2+x2^2-15
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [g,h,dg,dh] = exm1_ctr(x)
  g = [ 2*x(1)^2+x(2)^2-15 ]; % inequality constraints
  h = [];                     % equality constraints
  if nargout > 2
    dg = [ 4*x(1); 2*x(2); 0 ]; % gradient of inequality constraints
    dh = [];                    % gradient of equality constraints
  end
end
```

`fmincon` terminates normally, i.e., with `iout=1`. The iterates converge to the point $\mathbf{x}^\star \approx \begin{pmatrix} 2.7386 & 0 & 5.7386 \end{pmatrix}^\mathsf{T}$, and the corresponding optimal solution value is $f(\mathbf{x}^\star) \approx -77.7950$. Moreover, the active constraints at $\mathbf{x}^\star$ are g_1, g_2 and the lower bound on x_2.

(b) Repeat the numerical optimization from a different starting point, e.g., $\mathbf{x}^0 = \begin{pmatrix} 4 & 0 & 0 \end{pmatrix}^\mathsf{T}$. Does `fmincon` converge to the same solution point? Could this be expected?

Solution. The iterates converge to the same optimal solution point $\mathbf{x}^0 \approx \begin{pmatrix} 2.7386 & 0 & 5.7386 \end{pmatrix}^\mathsf{T}$. This could be expected since Problem (1) is strictly convex.

(c) Get the values $\boldsymbol{\nu}^\star$ of the Lagrange multipliers at $\mathbf{x}^\star$, as well as the gradients of the objective function and constraints. Check that the optimal solution point is (i) a regular point and (ii) a KKT point.

Solution. These checks can be done by modifying the file `exm1.m` as follows:

exm1.m

```
clear all
options = optimset('Display', 'iter', 'GradObj', 'on', ...
                   'GradConstr', 'on', 'DerivativeCheck', 'on', ...
                   'LargeScale', 'off', 'HessUpdate', 'bfgs', ...
                   'Diagnostics', 'on',  'TolX', 1e-7, ...
                   'TolFun', 1e-7, 'TolCon', 1e-7,
                   'MaxFunEval', 100, 'MaxIter', 100 )
x0 = [ 1; 1; 1 ];
xL = [ 0; 0; 0 ];
```

```
A  = [ -1 2 1 ];
b  = [ 3 ];

%Solve NLP Problem
[xopt, fopt, iout, output, mopt ] = fmincon( @exm1_fun, x0, A, b, [], [], xL, [], ...
                                              @exm1_ctr, options );
[fopt, dfopt] = exm1_fun(xopt);
[gopt, dum, dgopt, ddum] = exm1_ctr(xopt);

%Check regularity
rank = rank( [ dgopt(1)  A(1)   0;
               dgopt(2)  A(2)  -1;
               dgopt(3)  A(3)   0; ]);
disp(sprintf('Rank of Active Constraints: %d', rank))

%Check KKT conditions
for i= 1:3
  KKT(i) = dfopt(i) + mopt.ineqnonlin*dgopt(i) + mopt.ineqlin*A(i) - mopt.lower(i);
end
disp(sprintf('KKT Conditions = %g %g %g', KKT(1), KKT(2), KKT(3)))
```

The values of the Lagrange multipliers are $\nu_1 \approx 1.1432$ (constraint g_1), $\nu_2 = 12$ (constraint g_2), and $\mu_2^- \approx 24.7386$ (lower bound on x_2).

i. It is found that the rank of the Jacobian of the active constraints at $\mathbf{x}^\star$ is 3, hence $\mathbf{x}^\star$ is a regular point.

ii. Regarding KKT conditions at $\mathbf{x}^\star$, the constraints are all satisfied, the dual feasibility conditions are satisfied within 10^{-8}, and the Lagrange multipliers of the active inequality constraints are all nonnegative. So $\mathbf{x}^\star$ is a KKT point too.

(d) Consider the perturbed problem:

$$
\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^3}\ & f(\mathbf{x}) := x_1^2 + x_1x_2 + 2x_2^2 - 6x_1 - 2x_2 - 12x_3 \\
\text{s.t.}\ \ & g_1(\mathbf{x},\theta) := 2x_1^2 + x_2^2 \le \theta \\
& g_2(\mathbf{x}) := x_1 - 2x_2 - x_3 \ge 3 \\
& x_1, x_2, x_3 \ge 0,
\end{aligned}
\tag{3}
$$

where θ stands for the perturbation parameter. Solve (3) for N equally spaced values of θ in the range $[0, 30]$ (e.g., use a resolution of 0.5 for θ). Let us denote the optimal solution points as $\boldsymbol{\xi}^\star(\theta)$ and the optimal value of the Lagrange multiplier associated to the perturbed constraint as $\omega^\star(\theta)$.

Solution. A possible modification of the mfiles for solving the perturbed NLP problem is as follows:

exm1.m

```
clear all
options = optimset('Display', 'iter', 'GradObj', 'on', ...
                   'GradConstr', 'on', 'DerivativeCheck', 'on', ...
                   'LargeScale', 'off', 'HessUpdate', 'bfgs', ...
                   'Diagnostics', 'on',  'TolX', 1e-7, ...
                   'TolFun', 1e-7, 'TolCon', 1e-7,
                   'MaxFunEval', 100, 'MaxIter', 100 )
```

```
xO = [ 1; 1; 1 ];
xL = [ 0; 0; 0 ];
A  = [ -1 2 1 ];
b  = [ 3 ];

%Solve NLP Perturbed Problem
for ic = 1:61
  ci(ic) = (ic-1)/2.;
  [xopt, fopt, iout, output, mopt ] = fmincon( @exm1_fun, x0, A, b, [], [], xL, ...
                                               [], @(x)exm1_ctr(x,ci(ic)), options );
  fi(ic) = fopt;
  ni(ic) = mopt.ineqnonlin;
end

%Plot Results
clf;
figure(1);
plot( ci, fi, 'b' );
axis([ 0 30 0 10 ]);
figure(2);
plot( ci, ni, 'b' );
axis([ 0 30 0 10 ]);

%Estimated slope f(xopt) vs. ci
slope = (fi(32)-fi(30))/(ci(32)-ci(30));
disp(sprintf('Slope of f(xi(ci)) at ci=15: %d', slope))
```

exm1_ctr.m

```
%%%%%%%%%%%%%%%%%%%%%% CONSTRAINTS %%%%%%%%%%%%%%%%%%%%%
% nonlin. ineq. constraint: g1(x) := 2*x1^2+x2^2-15
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function [g,h,dg,dh] = exm1_ctr(x,c)
  g = [ 2*x(1)^2+x(2)^2-c ]; % inequality constraints
  h = [];                     % equality constraints
  if nargout > 2
    dg = [ 4*x(1); 2*x(2); 0 ]; % gradient of inequality constraints
    dh = [];                     % gradient of equality constraints
  end
end
```

i. Plot $f(\boldsymbol{\xi}^\star(\theta))$ versus θ, and estimate the slope of this curve at $\theta = 15$. What does the corresponding value represent?

Solution. The resulting curve is shown in the left plot of Fig. 1 below. An estimation of the slope of this curve at $\theta = 15$ is obtained via a (centered) finite difference formula as:

$$\frac{\partial f(\boldsymbol{\xi}^\star(\theta))}{\partial\theta} \approx \frac{f(\boldsymbol{\xi}^\star(15+\Delta\theta)) - f(\boldsymbol{\xi}^\star(15-\Delta\theta))}{2\Delta\theta} \approx -1.143396.$$

This value corresponds to minus the Lagrange multiplier relative to the inequality g_1,

$$\left.\frac{\partial f(\boldsymbol{\xi}^\star(\theta))}{\partial\theta}\right|_{\theta=15} = -\nu_1^\star \approx 1.1432.$$

Here, we can talk of *the* value of the Lagrange multiplier, for a regularity condition holds at the optimal solution point, which guarantees uniqueness of the Lagrange multipliers; moreover, (3) being strictly convex, $\boldsymbol{\xi}^\star(\theta)$ is a unique for each value of θ,

$$\{\boldsymbol{\xi}^\star(\theta)\} = \arg\min\{f(\mathbf{x}) : g_1(\mathbf{x};\theta) \le 0, g_2(\mathbf{x}) \le 0, x_1, x_2, x_3 \ge 0\}.$$

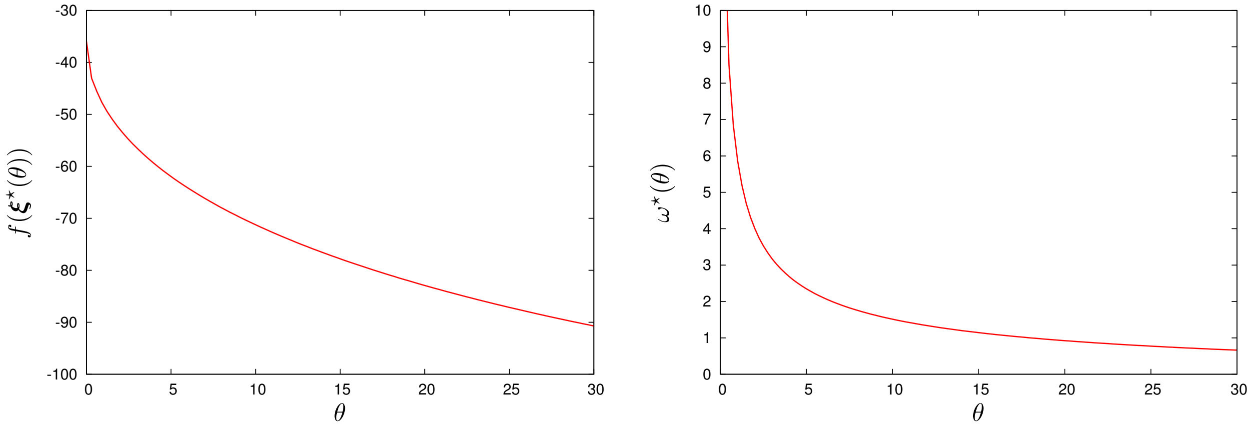

Figure 1: Plots of $f(\boldsymbol{\xi}^\star(\theta))$ vs θ (left plot) and $\omega^\star(\theta))$ vs θ (right plot) for Problem (3).

ii. Plot $\omega^\star(\theta)$ versus θ. Comment this plot and, in particular, explain the behavior at $\theta = 0$. What is be the slope of the curve $f(\boldsymbol{\xi}^\star(\theta))$ versus θ at $\theta = 0$?

Solution. At $\theta = 0$, it is easily seen that the only feasible points for the problem are points on the line segment $I := \{\mathbf{x} \in \mathbb{R}^3 : x_1 = x_2 = 0, 0 \le x_3 \le 3\}$. Since the objective function is inversely proportional to x_3, the optimal solution is obviously $\boldsymbol{\xi}^\star(0) = \begin{pmatrix} 0 & 0 & 3 \end{pmatrix}^\mathsf{T}$. The active constraints at $\boldsymbol{\xi}^\star(0)$ are g_1, g_2 as well as the lower bounds $x_1 \ge 0$, $x_2 \ge 0$. 4 constraints are thus active, while the search space is of dimension 3 only, i.e., the gradients of the active constraints at $\boldsymbol{\xi}^\star(0)$ cannot be linearly independent and $\boldsymbol{\xi}^\star(0)$ is *not* a regular point. This, in turn, indicates that the KKT conditions may *not* hold at $\boldsymbol{\xi}^\star(0)$, even though $\boldsymbol{\xi}^\star(0)$ is a local minimum point. The non-existence of a Lagrange multiplier at $\theta = 0$ is confirmed from the right plot of Fig. 1; it is seen that the value of the Lagrange multiplier goes to $+\infty$ as $\theta \to 0+$.
Finally, since $\frac{\partial f(\boldsymbol{\xi}^\star(\theta))}{\partial \theta} = -\omega^\star(\theta)$, for each $\theta > 0$, then

$$\lim_{\theta \to 0^+} \frac{\partial f(\boldsymbol{\xi}^\star(\theta))}{\partial \theta} = -\infty.$$

This behavior can be seen from the left plot of Fig. 1.