Barrier (or Interior Point) Method for Linear Programming

Consider the following Linear Programming problem:

$$
\begin{aligned}
(P) \quad \min \quad & \underline{c}^\mathsf{T} \underline{x} && (1)\\
\text{s.t.} \quad & A\underline{x} = \underline{b} && (2)\\
& \underline{x} \geq \underline{0}. && (3)
\end{aligned}
$$

The constraints that make the problem (P) "hard" are the variable nonnegativity constraints. We consider applying the logarithmic barrier method (or interior point method) to the problem where only the constraints $\underline{x} \geq \underline{0}$ are relaxed. Let $(P(\mu))$ be the logarithmic barrier problem obtained for a given value of the barrier parameter μ, that is:

$$
\begin{aligned}
(P(\mu)) \quad \min \quad & \underline{c}^\mathsf{T} \underline{x} - \mu \sum_{i=1}^{n} \ln x_i && (4)\\
\text{s.t.} \quad & A\underline{x} = \underline{b}. && (5)
\end{aligned}
$$

a) Determine the Karush-Kuhn-Tucker conditions for the original problem (P). Are they necessary and/or sufficient for the optimality in any point of the feasible region?

Consider the KKT conditions of the problem (P). If we perturb each complementarity condition by a same parameter $\mu > 0$, we obtain the following perturbed conditions:

$$
\begin{aligned}
\underline{c} - \underline{v} + A^\mathsf{T} \underline{u} &= \underline{0} && (6)\\
A\underline{x} &= \underline{b} && (7)\\
\underline{x} &\geq \underline{0} && (8)\\
\underline{v} &\geq \underline{0} && (9)\\
x_j v_j &= \mu \qquad \text{for } j = 1, \ldots, n. && (10)
\end{aligned}
$$

Definition: Let $(\underline{x}(\mu), \underline{u}(\mu), \underline{v}(\mu))$ be vectors satisfying the conditions (6)-(10). As $\mu > 0$ changes, the triplet of vectors $(\underline{x}(\mu), \underline{u}(\mu), \underline{v}(\mu))$ describe a trajectory called the *central path*. Observe that, when μ tends to 0, the complementary conditions tends to be satisfied.

More precisely, we can prove that, if $\mu \to 0$, the triplet $(\underline{x}(\mu), \underline{u}(\mu), \underline{v}(\mu))$ converges to a triplet $(\underline{x}^*, \underline{u}^*, \underline{v}^*)$ satisfying the KKT conditions of problem (P).

It is possible to verify that

$$
L(\underline{x}(\mu), \underline{u}(\mu), \underline{v}(\mu)) = \underline{c}^\mathsf{T} \underline{x}(\mu) - n\mu \leq \underline{c}^\mathsf{T} \underline{x}^* \leq \underline{c}^\mathsf{T} \underline{x}(\mu)
$$

where $\underline{x}^*$ is an optimal solution for (P) (for the proof, see Appendix 1). Thus, along the central path, for any $\mu > 0$ the value of the objective function of (P) $\underline{c}^\mathsf{T} \underline{x}(\mu)$ is at most $n\mu$ from the optimal value $\underline{c}^\mathsf{T} \underline{x}^*$.

b) Consider now the problem with logarithmic barrier $(P(\mu))$ for a given $\mu > 0$. Determine the Karush-Kuhn-Tucker conditions for such problem. Are they necessary and/or sufficient for optimality in any point of the feasible region?

c) Let $\underline{x}^*(\mu)$ be a solution satisfying the KKT conditions of $(P(\mu))$ for given multipliers $\underline{u}^*(\mu)$. Can the pair $(\underline{x}^*(\mu), \underline{u}^*(\mu))$ be part of a triplet $(\underline{x}^*(\mu), \underline{u}^*(\mu), \underline{v}^*(\mu))$ belonging to the central path?

d) Consider an iterative algorithm where we apply the Newton method for the solution of the barrier problem at each iteration. Show how one can adapt the rule to derive the Netwon step in order to satisfy a set of linear equality constraints.

e) Implement the above-method logarithmic barrier method in Matlab and solve the following problem:

$$
\begin{aligned}
\min \quad & x_1 - x_2 && (11)\\
\text{s.t.} \quad & -x_1 + x_2 \leq 1 && (12)\\
& x_1 + x_2 \leq 3 && (13)\\
& x_1, x_2 \geq 0. && (14)
\end{aligned}
$$

, starting from the feasible solution $\underline{\bar{x}} = (1, 1)$.

f) Solve the problem (11)-(14) with the simplex method, using AMPL and CPLEX. Verify the solution found with the interior point method is not on a vertex of the feasible polyhedron, as opposed to the one found by CPLEX. What is the reason?

g) Solve the problem obtained by replacing the objective function with $\frac{1}{2}x_1 - x_2$. Verify the solutions of the two methods coincide. Why?

A sketch of the function implementing the interior point method is in `ipm_stub.m`:

```
% interior point method for LPs of the form min c'*x : A*x <= b

function [xstar, vstar] = ipm(c, A, b, mu, x_init, gamma, epsilon)

  OPTIONS = [];
  [m, n] = size(A);

  %%bring to standard form
  s = b - A*x_init
  x = [ x_init ; s ]

  A = [A eye(m,m)]
  c = [c zeros(1,m)]
  %[m, n] = size(AA)



  grad = zeros(m+n, 1);
  H = zeros(m+n, m+n);
  d = zeros(m+n, 1);
  z = zeros(m, 1);

  xks = [x_init']

  %% iterative method
  while n * mu >= epsilon
    %% compute gradient and Hessian


    %% compute direction d with adapted Newton update


    alpha = fminbnd(@(alpha) % write here the objective function f(x+alpha*d), ...
            0, 1, OPTIONS);
    xstar = x + alpha * d;
    mu = gamma * mu;
    x = xstar;
    xks = [xks; x(1:n)']
  end


polyhedron_print(A,b); hold on;
plot(xks(:,1), xks(:,2), 'r.');

end % end of function
```

To represent the feasible region and the sequence of solutions, we include the function `polyhedron_print.m`:

```
function polyhedron_print(A, b)

  [m, n] = size(A);
  Ineq = A(:, 1 : n-m);
  rel = '';
  for i = 1 : m
    rel = [ rel '<' ];
  end
  t = extrpts(Ineq,rel,b);
  t = t(1:2,:);
  t = delcols(t);
  t1 = t(1,:);
  t2 = t(2,:);
  z = convhull(t1,t2);
  hold on
  patch(t1(z),t2(z),[0.9, 0.9, 0.5])

end % end of function
```

that makes use of the scriptst `vr.m`

```
function e = vr(m,i)

% The ith coordinate vector e in the m-dimensional Euclidean space.

e = zeros(m,1);
e(i) = 1;
```

`delcols.m`

```
function d = delcols(d)

% Delete duplicated columns of the matrix d.

d = union(d',d','rows')';
n = size(d,2);
j = [];
for k =1:n
   c = d(:,k);
   for l=k+1:n
      if norm(c - d(:,l),'inf') <= 100*eps
         j = [j l];
      end
   end
end
if ~isempty(j)
   j = sort(j);
   d(:,j) = [];
end
```

and `extrpts.m`

```
function vert = extrpts(A, rel, b)

% Extreme points vert of the polyhedral set

%               X = {x: Ax <= b or Ax >= b, x >= 0}.

% Inequality signs are stored in the string rel, e.g.,
% rel = '<<>' stands for <= , <= , and >= , respectively.

[m, n] = size(A);
nlv = n;
for i=1:m
   if(rel(i) == '>')
      A = [A -vr(m,i)];
   else
      A = [A vr(m,i)];
   end
   if b(i) < 0
      A(i,:) = - A(i,:);
      b(i) = -b(i);
   end
end
warning off
[m, n]= size(A);
b = b(:);
vert = [];
if (n >= m)
   t = nchoosek(1:n,m);
   nv = nchoosek(n,m);
   for i=1:nv
      y = zeros(n,1);
      x = A(:,t(i,:))\b;
      if all(x >= 0 & (x ~= inf & x ~= -inf))
         y(t(i,:)) = x;
         vert = [vert y];
      end
   end
else
   error('Number of equations is greater than the number of variables')
end
vert = delcols(vert);
vert = vert(1:nlv,:);
```

Appendix 1: value of the Lagrangian of (P) along the central path

Let $\left(\underline{\tilde{x}}\left(\mu\right), \underline{\tilde{u}}\left(\mu\right), \underline{\tilde{v}}\left(\mu\right)\right)$ be a solution on the central path. The Lagrangian of (P) is:

$$L\left(\underline{\tilde{x}}\left(\mu\right), \underline{\tilde{u}}\left(\mu\right), \underline{\tilde{v}}\left(\mu\right)\right) = \underline{c}^{\mathsf{T}}\underline{\tilde{x}} - \sum_{j=1}^{n} \tilde{v}_j \tilde{x}_j + \underline{\tilde{u}}^{\mathsf{T}}(A\underline{\tilde{x}} - \underline{b}) = \underline{c}^{\mathsf{T}}\underline{\tilde{x}} - \sum_{j=1}^{n} \tilde{v}_j \tilde{x}_j = \underline{c}^{\mathsf{T}}\underline{\tilde{x}} - n\mu,$$

since (6)-(10) are satisfied.

We can also write:

$$L\left(\underline{\tilde{x}}\left(\mu\right), \underline{\tilde{u}}\left(\mu\right), \underline{\tilde{v}}\left(\mu\right)\right) = c^{\mathsf{T}}\underline{\tilde{x}} - \sum_{j=1}^{n} \tilde{v}_j \tilde{x}_j = \left(-A^{\mathsf{T}}\underline{\tilde{u}} + \underline{\tilde{v}}^{\mathsf{T}}\right)\underline{x} - \sum_{j=1}^{n} \tilde{v}_j \tilde{x}_j = (-A^{\mathsf{T}}\underline{\tilde{u}})^{\mathsf{T}}\underline{\tilde{x}} = -\underline{\tilde{u}}^{\mathsf{T}}A\underline{\tilde{x}} = -\underline{\tilde{u}}^{\mathsf{T}}\underline{b}$$

The Lagrangian of (P) evaluated at a solution on the central path coincides with the value of the objective function of the dual problem.
[The dual of (P) is as follows:

$$\begin{aligned}(D) \quad & \max\ \underline{b}^{\mathsf{T}}\underline{u} \\ & \text{s.t. } A^{\mathsf{T}}\underline{u} \le \underline{c}.\end{aligned}$$

Since the variables $\underline{u}$ are iunrestricted in sign, the following problem has the same optimal value of (D):

$$\begin{aligned}\max\ & -\underline{b}^{\mathsf{T}}\underline{u} \\ \text{s.t. } & -A^{\mathsf{T}}\underline{u} \le \underline{c}.\end{aligned}$$

]

Denoting by $\underline{x}^*$ an optimal solution for (P), it follows that

$$c^{\mathsf{T}}\underline{\tilde{x}} - n\mu = -\underline{\tilde{u}}^{\mathsf{T}}\underline{b} \le c^{\mathsf{T}}\underline{x}^* \qquad \text{(due to weak duality in linear programming)}$$
$$c^{\mathsf{T}}\underline{x}^* \le c^{\mathsf{T}}\underline{\tilde{x}} \qquad \text{(since } c^{\mathsf{T}}\underline{x}^* \text{ is the optimal value),}$$

implying that

$$c^{\mathsf{T}}\underline{\tilde{x}} - n\mu \le c^{\mathsf{T}}\underline{x}^* \le c^{\mathsf{T}}\underline{\tilde{x}}$$

Then, along the central path, when μ tends to 0, the Lagrangian of (P) tends indeed to the optimal vlaue of the objective function.