1) Consider the constrained optimization problem (P)

$$\begin{cases} \min\ x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_2x_3 - 3x_1 - 4x_2 - 5x_3 \\ 2x_1 + x_2 + x_3 \leq 20 \\ x_1 \geq 2 \\ x_2 \leq 3 \\ x_3 \geq 4 \\ x \in \mathbb{R}^3 \end{cases}$$

(a) Is the problem (P) convex?

(b) Does (P) admit a global optimal solution?

(c) Apply the logarithmic barrier method with starting point $x^0 = (3, 2, 5)$, $\varepsilon^0 = 1$, $\tau = 0.5$ and tolerance 10^{-3}. How many iterations are needed by the algorithm? Write the vector x found at the last three iterations.

(d) Is the obtained solution a global minimum of the given problem? Justify the answer.

SOLUTION

(a)-(b) The problem is convex since the objective function is a quadratic strongly convex function $0.5x^TQx + c^Tx$ (the eigenvalues of the Hessian Q are: 0.5858 2.0000, 3.4142 and are evaluated in the Matlab solution) and the constraints $Ax - b \leq 0$ are linear. This also implies that (P) admits a unique global optimal solution. See the Matlab solution for the definition of the matrices Q and A and the vectors c and b.

(c) **Matlab solution**

```
global Q c A b eps;                                    %% data

Q = [ 2 -1 0 ;-1 2 1; 0 1 2 ] ;
c = [ -3 ; -4; -5 ] ;
A = [2 1 1;-1 0 0; 0 1 0 ;0 0 -1 ];
b = [ 20; -2 ; 3; -4  ];

eig(Q)

delta = 1e-3 ;
tau = 0.5 ;
eps1 = 1 ;
x0 = [3;2;5];


x = x0;                                          %% barrier method
eps = eps1 ;
m = size(A,1) ;

SOL=[]

while true
    [x,pval] = fminunc(@logbar,x);
    gap = m*eps;
    SOL=[SOL;eps,x',gap,pval];
    if gap < delta
        break
    else
        eps = eps*tau;
    end
end

fprintf('\t eps \t x(1) \t x(2) \t x(3) \t gap \t  pval \n\n');

SOL

function v = logbar(x)                                    %% logarithmic barrier function
```

```
    global Q c A b eps

    v = 0.5*x'*Q*x + c'*x ;

    for i = 1 : length(b)
        v = v - eps*log(b(i)-A(i,:)*x) ;
    end

end
```

We obtain the following solution:

```
 iter    eps       x(1)      x(2)     x(3)       gap       pval

SOL =

11.    0.0010    2.0253    1.0122    4.0002    0.0039   -6.9898
12.    0.0005    2.0190    1.0124    4.0001    0.0020   -6.9944
13.    0.0002    2.0190    1.0124    4.0001    0.0010   -6.9969
```

The iterations of the algorithm are 13 and the found solution is $x^* \approx (2, 1, 4)$ with optimal value $val(P) \approx -7$.

(d) The point $x^* = (2, 1, 4)$ is a global minimum. The algorithm converges to a solution of the KKT conditions associated with (P), that in the present case are sufficient for optimality. The KKT conditions for (P) are given by

$$\begin{cases} Qx + c + A^T\lambda = 0 \\ \lambda^T(Ax - b) = 0 \\ Ax - b \leq 0, \\ \lambda \geq 0 \\ \lambda \in \mathbb{R}^4, x \in \mathbb{R}^3 \end{cases}$$

It is easy to verify that $x^* = (2, 1, 4)$ with $\lambda^* = (0, 0, 0, 4)$ is a solution of the previous system.

2) Consider a clustering problem for the following set of 18 patterns given by the columns of the matrix:

$$D = \begin{pmatrix} 1 & 2 & 1 & 4.8 & 2.7 & 4.8 & 4.0 & 3.7 & 3.5 & 3.2 & 1.8 & 2.4 & 2.6 & 6 & 5.2 & 7.7 & 8 & 6 \\ 6 & 8 & 7 & 7.3 & 6.3 & 8 & 6.7 & 5.1 & 7.8 & 5.7 & 6.5 & 6.2 & 7.2 & 8 & 7.5 & 7 & 4 & 5 \end{pmatrix}$$

(a) Write the optimization model with 2-norm and $k = 3$ clusters;

(b) Solve the problem using the k-means algorithm with $k = 3$ and starting from centroids $x^1 = (1, 3), x^2 = (2, 4), x^3 = (4, 1)$. Write explicitly the vector of the obtained centroids and the value of the objective function;

(c) Solve the problem using the k-means algorithm with $k = 3$ and starting from centroids $x^1 = (6, 3), x^2 = (4, 2), x^3 = (7, 4)$. Write explicitly the vector of the obtained centroids, the value of the objective function and the clusters;

(d) Is any of the solutions found at points (b) and (c) a global optimum? Justify the answer.

SOLUTION

(a) Let $\ell = 18$, $(p^i)^T$ be the i-th row of the matrix D^T, $i = 1, .., 18$, the problem can be formulated by

$$\begin{cases} \min\limits_{x} \sum\limits_{i=1}^{\ell} \min\limits_{j=1,2,3} \|p^i - x^j\|_2^2 \\ x^j \in \mathbb{R}^2 \qquad \forall\ j = 1, 2, 3 \end{cases} \tag{1}$$

which is equivalent to

$$\begin{cases} \min\limits_{x,\alpha}\ f(x, \alpha) := \sum\limits_{i=1}^{\ell} \sum\limits_{j=1}^{3} \alpha_{ij} \|p^i - x^j\|_2^2 \\ \sum\limits_{j=1}^{3} \alpha_{ij} = 1 \quad \forall\ i = 1, \dots, \ell \\ \alpha_{ij} \geq 0 \qquad \forall\ i = 1, \dots, \ell,\ j = 1, 2, 3 \\ x^j \in \mathbb{R}^2 \qquad \forall\ j = 1, 2, 3. \end{cases} \tag{2}$$

provided that we look for an optimal solution with $\alpha \in \{0, 1\}^{\ell \times k}$, which can be done by the k-means algorithm.

(b)-(d) **Matlab solution**

```
data = [ 1    6
         2     8
         1    7
         4.8    7.3
         2.7    6.3
         4.8    8
         4.0    6.7
         3.7    5.1
         3.5    7.8
         3.2    5.7
         1.8    6.5
         2.4    6.2
         2.6    7.2
          6     8
         5.2    7.5
         7.7    7.0
          8      4
          6      5
                    ];

l = size(data,1); % number of patterns

k=3;

InitialCentroids=[1,3;2,4;4,1];
 % InitialCentroids=[6,3;4,2;7,4];

[x,cluster,v] = kmeans1(data,k,InitialCentroids)
```

```
function [x,cluster,v] = kmeans1(data,k,InitialCentroids)

l = size(data,1); % number of patterns

x = InitialCentroids;

cluster = zeros(l,1);                          % initialize clusters
for i = 1 : l
    d = inf;
    for j = 1 : k
        if norm(data(i,:)-x(j,:)) < d
            d = norm(data(i,:)-x(j,:));
            cluster(i) = j;
        end
    end
end

vold = 0;                                 % compute the objective function value
for i = 1 : l
    vold = vold + norm(data(i,:)-x(cluster(i),:))^2 ;
end

while true
          for j = 1 : k                                      % update centroids
        ind = find(cluster == j);
        if isempty(ind)==0
            x(j,:) = mean(data(ind,:),1);
        end
    end


    for i = 1 : l                            % update clusters
        d = inf;
        for j = 1 : k
            if norm(data(i,:)-x(j,:)) < d
                d = norm(data(i,:)-x(j,:));
                cluster(i) = j;
            end
        end
    end

    v = 0;                                % update objective function
    for i = 1 : l
        v = v + norm(data(i,:)-x(cluster(i),:))^2 ;
    end

    if vold - v < 1e-5                       % stopping criterion
        break
    else
        vold = v;
    end
end
end
```

In case (b) we obtain the solution:

```
x =

    1.0000    3.0000     % centroids
    3.0500    6.8071
    6.9250    6.0000

v =    47.8518       %value of the objective function
```

In case (c) we obtain the solution:

```
x =

    3.6333    5.8333
    2.1250    6.8750
    6.0714    6.6857


cluster = 2 2 2 3 2 3 1 1 2  1  2  2  2 3 3 3 3 3  % The j-th component is the cluster assigned to pattern j

v =    35.7662
```

(d) It is possible to improve the previous solutions, with a multistart approach. For example, starting from centroids $x^1 = (1,4), x^2 = (3,4), x^3 = (4,2)$, we obtain the solution

```
x =
    2.2667    6.4444
    4.7167    7.5500
    7.2333    5.3333

v=  24.8189

cluster=   1 1 1 2 1 2 2  1 2  1 1 1 1 2 2 3 3 3
```

3) Consider the following multiobjective optimization problem (P):

$$\begin{cases} \min\ (x_1^2 - x_2^2,\ -x_1) \\ x_1 + x_2^2 \le 0 \end{cases}$$

(a) Is the given problem (P) convex?

(b) Prove that (P) admits a Pareto minimum point.

(c) Find a suitable subset of Pareto minima by means of the scalarization method.

SOLUTION

(a) The problem is not convex since $f_1(x_1, x_2) := x_1^2 - x_2^2$ is not convex;

(b) We preliminarly prove that the scalarized problem (P_{α_1}) defined by

$$\begin{cases} \min\ \alpha_1(x_1^2 - x_2^2) + (1-\alpha_1)(-x_1) =: \psi_{\alpha_1}(x) \\ x_1 + x_2^2 \le 0 \end{cases}$$

admits an optimal solution for every $0 \le \alpha_1 \le 1$.

To this aim consider the 0- sublevel set of the function $\psi_{\alpha_1}(x)$ on the feasible set:

$$\begin{cases} \alpha_1(x_1^2 - x_2^2) + (1-\alpha_1)(-x_1) \le 0 \\ x_1 + x_2^2 \le 0 \end{cases}$$

By the second inequality it follows

$$-\alpha_1 x_2^2 \ge \alpha_1 x_1, \quad \forall \alpha_1 \ge 0$$

By the first inequality we have:

$$\alpha_1 x_1^2 + \alpha_1 x_1 + (1-\alpha_1)(-x_1) \le \alpha_1 x_1^2 - \alpha_1 x_2^2 + (1-\alpha_1)(-x_1) \le 0$$

Then

$$\alpha_1 x_1^2 - x_1(1 - 2\alpha_1) \le 0.$$

The solutions of the previous inequality are:

i) $0 \le x_1 \le \frac{1-2\alpha_1}{\alpha_1}$ for $0 < \alpha_1 < \frac{1}{2}$;

ii) $\frac{1-2\alpha_1}{\alpha_1} \le x_1 \le 0$ for $\frac{1}{2} \le \alpha_1 \le 1$;

Being $x_2^2 \le -x_1$, in case (i) we have $x_2^2 \le -x_1 \le 0$, i.e. $x_1 = x_2 = 0$, while in case ii) we have $x_2^2 \le -x_1 \le \frac{2\alpha_1 - 1}{\alpha_1}$ which implies

$$-\sqrt{\frac{2\alpha_1 - 1}{\alpha_1}} \le x_2 \le \sqrt{\frac{2\alpha_1 - 1}{\alpha_1}}, \quad \frac{1}{2} \le \alpha_1 \le 1.$$

In both cases, the 0-level set of $\psi_{\alpha_1}(x)$ is bounded. Similarly, it can be proved that it is bounded for $\alpha_1 = 0$. This implies that the scalarized problem (P_{α_1}) admits a global obtimal solution for every $0 \le \alpha_1 \le 1$, which is a Pareto minimum point, for $0 < \alpha_1 < 1$.

(c) Let us find Pareto minima of (P) by solving the KKT conditions for (P_{α_1}), for $0 < \alpha_1 < 1$. (P_{α_1}) is a differentiable problem and fulfils the Abadie constraints qualifications, being the gradient of the constraint function $\nabla g(x) = (1, 2x_2) \ne (0, 0)$. The KKT system provides a necessary condition for an optimal solution of (P_{α_1}), that we have proved to exist.

$$\begin{cases} 2\alpha_1 x_1 - (1-\alpha_1) + \lambda = 0 \\ -2\alpha_1 x_2 + 2\lambda x_2 = 0 \\ \lambda(x_1 + x_2^2) = 0 \\ x_1 + x_2^2 \le 0 \\ \lambda \ge 0 \\ 0 < \alpha_1 < 1, \end{cases}$$

or, equivalently

$$\begin{cases} 2\alpha_1 x_1 + \alpha_1 - 1 + \lambda = 0 \\ 2x_2(\lambda - \alpha_1) = 0 \\ \lambda(x_1 + x_2^2) = 0 \\ x_1 + x_2^2 \le 0 \\ \lambda \ge 0 \\ 0 < \alpha_1 < 1, \end{cases} \tag{3}$$

By the second equation, we have the following cases: $I)\ x_2 = 0,\ II)\ \lambda - \alpha_1 = 0$.

In case I), the system (3) becomes

$$\begin{cases} 2\alpha_1 x_1 + \alpha_1 - 1 + \lambda = 0 \\ x_2 = 0 \\ \lambda x_1 = 0 \\ x_1 \leq 0 \\ \lambda \geq 0 \\ 0 < \alpha_1 < 1, \end{cases}$$

and admits the solutions:

$$\begin{cases} x_1 = x_2 = 0 \\ \lambda = 1 - \alpha_1 \\ 0 < \alpha_1 < 1, \end{cases}$$

In case II), the system (3) becomes

$$\begin{cases} 2\alpha_1 x_1 + 2\alpha_1 - 1 = 0 \\ \lambda = \alpha_1 \\ \alpha_1(x_1 + x_2^2) = 0 \\ x_1 + x_2^2 \leq 0 \\ 0 < \alpha_1 < 1, \end{cases}$$

or, equivalently,

$$\begin{cases} x_1 = \frac{1-2\alpha_1}{2\alpha_1} \\ \lambda = \alpha_1 \\ x_1 = -x_2^2 \\ x_1 + x_2^2 \leq 0 \\ 0 < \alpha_1 < 1, \end{cases}$$

which admits the solutions:

$$\begin{cases} x_1 = \frac{1-2\alpha_1}{2\alpha_1} \\ x_2 = \pm\sqrt{\frac{2\alpha_1-1}{2\alpha_1}} \\ \lambda = \alpha_1 \\ \frac{1}{2} \leq \alpha_1 < 1, \end{cases}$$

Both obtained points provide the same value of the function $\psi_{\alpha_1}(x)$, so they are both optimal solutions of P_{α_1}.

In conclusion, the points

$$\{(x_1, x_2) = \left(\frac{1-2\alpha_1}{2\alpha_1}, \pm\sqrt{\frac{2\alpha_1 - 1}{2\alpha_1}}\right),\ \frac{1}{2} \leq \alpha_1 < 1\}$$

are all Pareto minima for (P).

4) Consider the following bimatrix game:

$$C_1 = \begin{pmatrix} 2 & -1 & -1 \\ 0 & 2 & 1 \\ 1 & 3 & 4 \end{pmatrix} \qquad C_2 = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & -1 \\ 4 & 0 & 1 \end{pmatrix}$$

(a) Find the strictly dominated strategies and reduce the matrices accordingly.

(b) Find the set of pure strategies Nash equilibria, if any. Alternatively, show that no pure strategies Nash equilibrium exists.

(c) Find a mixed strategies Nash equilibrium.

SOLUTION

(a) Strategy 3 of Player 1 is dominated by Strategy 2, so that row 3 in the two matrices can be deleted. We obtain:

$$C_1^R = \begin{pmatrix} 2 & -1 & -1 \\ 0 & 2 & 1 \end{pmatrix} \qquad C_2^R = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & -1 \end{pmatrix}$$

Now Strategy 2 of Player 2 is dominated by Strrtegy 1 so that column 2 in the two matrices can be deleted. The reduced game is given by the matrices

$$C_1^R = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} \qquad C_2^R = \begin{pmatrix} 1 & 4 \\ 2 & -1 \end{pmatrix}$$

For Player 1, the possible couples of pure strategies Nash equilibria could be $(2,1)$ and $(1,3)$ (corresponding to minima on the columns of C_1^R, while considering Player 2 the pure strategies Nash equilibria could be $(1,1)$ and $(2,3)$ (corresponding to minima on the rows of C_2^R). Since there are no common couples, no pure strategies Nash equilibria exist.

This will also be shown in part (b) in the wider context of mixed strategies Nash equilibria.

(b) Consider the reduced game obtained in (a). The optimization problem associated with Player 1 is

$$\begin{cases} \min\ x^T C_1^R y = (2x_1)y_1 + (-x_1 + x_2)y_3 \\ x_1 + x_2 = 1 \\ x_1, x_2 \geq 0 \end{cases} \equiv \begin{cases} \min\ (4y_1 - 2)x_1 + 1 - y_1 \\ 0 \leq x_1 \leq 1 \end{cases} \qquad (P_1(y_1))$$

Then, the best response mapping associated with $P_1(y_1)$ is:

$$B_1(y_1) = \begin{cases} 0 & \text{if } y_1 \in (1/2, 1] \\ [0,1] & \text{if } y_1 = 1/2 \\ 1 & \text{if } y_1 \in [0, 1/2) \end{cases}$$

Similarly, the optimization problem associated with Player 2 is

$$\begin{cases} \min\ x^T C_2^R y = (x_1 + 2x_2)y_1 + (4x_1 - x_2)y_3 \\ y_1 + y_3 = 1 \\ y_1, y_3 \geq 0 \end{cases} \equiv \begin{cases} \min\ (3 - 6x_1)y_1 + 5x_1 - 1 \\ 0 \leq y_1 \leq 1 \end{cases} \qquad (P_2(x_1))$$

Then, the best response mapping associated with $P_2(x_1)$ is:

$$B_2(x_1) = \begin{cases} 0 & \text{if } x_1 \in [0, 1/2) \\ [0,1] & \text{if } x_1 = 1/2 \\ 1 & \text{if } x_1 \in (1/2, 1] \end{cases}$$

The only couple (x_1, y_1) such that $x_1 \in B_1(y_1)$ and $y_1 \in B_2(x_1)$ are $x_1 = \frac{1}{2}, y_1 = \frac{1}{2}$, so that, recalling that $x_3 = 0$ and $y_2 = 0$,

$$(x_1, x_2, x_3) = (\frac{1}{2}, \frac{1}{2}, 0), \quad (y_1, y_2, y_3) = (\frac{1}{2}, 0, \frac{1}{2}),$$

is a mixed strategies Nash equilibrium and no pure strategies Nash equilibrium exists.