1) Consider the unconstrained optimization problem

$$\begin{cases} \min\ 2x_1^2 + x_2^2 + 2x_3^2 + 4x_4^2 + x_1x_2 + 2x_1x_4 + 2x_3x_4 - 2x_1 + 8x_2 + 6x_3 + 4x_4 \\ x \in \mathbb{R}^4 \end{cases}$$

(a) Apply the gradient method with exact line search, with starting point $x^0 = (1, 0, 0, 1)$ and using $\|\nabla f(x)\| < 10^{-3}$ as stopping criterion. How many iterations are needed by the algorithm? Write the vector found at the last three iterations.

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

SOLUTION

(a) The objective function $f(x)$ is quadratic, i.e., $f(x) = (1/2)x^TQx + c^Tx$ with

$$Q = \begin{pmatrix} 4 & 1 & 0 & 2 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 4 & 2 \\ 2 & 0 & 2 & 8 \end{pmatrix} \qquad c^T = (-2, 8, 6, 4)$$

Matlab solution

```
Q = [4 1 0 2;1 2 0 0;0 0 4 2;2 0 2 8];
c = [-2 8 6 4]';

disp('Eigenvalues of Q:')
eig(Q)

tolerance = 10^(-3);    % parameters

 x = [1 0 0 1]';     % starting point

X=[ ];

for ITER=1:1000
    v = 0.5*x'*Q*x + c'*x;
    g = Q*x + c ;
    X=[X;ITER,x',v,norm(g)];

    % stopping criterion
    if norm(g) < Tolerance
        break
    end

    %   search direction
    d = -g;

    %   exact line search
    t = norm(g)^2/(d'*Q*d) ;

    %   new point
    x = x + t*d ;
end
disp('optimal solution')
x
disp('optimal value')
v
disp('gradient norm at the solution')
norm(g)
disp('number of iterations')
ITER
```

%%%%%%%%%%%%%%%%%%%%

We obtain the following solution:

```
Eigenvalues of Q:


ans =

    1.4341
    2.8417
    4.2457
    9.4785

>> ITER

ITER =

    26

>> X(24:26,:)

24.0000    2.1459   -5.0724   -1.1221   -0.7560  -27.3171    0.0015
25.000      2.1461   -5.0727   -1.1220   -0.7559  -27.3171    0.0015
26.0000    2.1461   -5.0728   -1.1220   -0.7560  -27.3171    0.0007
```

The iterations of the algorithm are 26. In particular, the gradient norm evaluated at the final point is: 0.0007.

(b) The found point $x = (2.1461 - 5.0728 - 1.1220 - 0.7560)$ is an approximation of the global minimum which exists since the objective function is strongly convex: in fact the eigenvalues of the Hessian of f are all strictly positive. The analytic expression of the global minimum is $\bar{x} = -Q^{-1}c$, by Matlab we find

```
x=-inv(Q)*c

ans =

    2.1463
   -5.0732
   -1.1220
   -0.7561
```

2) Consider a binary classification problem with the data sets A and B given by the row vectors of the matrices:

$$A = \begin{pmatrix} 8 & 1.5 \\ 6 & 2 \\ 7 & 0.5 \\ 2.76 & 0.46 \\ 0.97 & 8.23 \\ 9.5 & 0.34 \\ 4.38 & 3.81 \\ 1.86 & 4.89 \\ 2.75 & 6.75 \\ 6.5 & 1.5 \end{pmatrix}, \quad B = \begin{pmatrix} 10 & 2.4 \\ 5 & 2 \\ 7.51 & 2.55 \\ 5.05 & 7 \\ 9 & 9.6 \\ 8.40 & 2.54 \\ 8.14 & 2.43 \\ 9.3 & 3.45 \\ 6.16 & 5 \\ 3.5 & 8.5 \end{pmatrix}$$

(a) Write the linear SVM model with soft margin to find the separating hyperplane;

(b) Solve the dual problem with parameter $C = 30$ and find the optimal hyperplane. Write explicitly the vector of the optimal solution of the dual problem;

(c) Find the misclassified points of the data sets A and B by means of the dual solution;

(d) Classify the new point $(4, 3)$.

SOLUTION

(a) Let $(x^i)^T$ be the i-th row of the matrix A, for $i = 1, .., 10$ and of the matrix B for $i = 11, .., 20$. For any point x^i, define the label:

$$y^i = \begin{cases} 1 & \text{if } i = 1, .., 10 \\ -1 & \text{if } i = 11, .., 20 \end{cases}$$

The formulation of the linear SVM with soft margin is

$$\begin{cases} \min\limits_{w,b,\xi} \ \frac{1}{2}\|w\|^2 + 30 \sum\limits_{i=1}^{20} \xi_i \\ 1 - y^i(w^T x^i + b) \leq \xi_i, \qquad \forall\ i = 1, \dots, 20 \\ \xi_i \geq 0, \qquad \forall\ i = 1, \dots, 20 \end{cases} \tag{1}$$

(b) The dual problem of (1) is

$$\begin{cases} \max\limits_{\lambda} \ -\frac{1}{2} \sum\limits_{i=1}^{20} \sum\limits_{j=1}^{20} y^i y^j (x^i)^T x^j \lambda_i \lambda_j + \sum\limits_{i=1}^{20} \lambda_i \\ \sum\limits_{i=1}^{20} \lambda_i y^i = 0 \\ 0 \leq \lambda_i \leq 30 \qquad i = 1, \dots, 20 \end{cases}$$

Matlab solution

```
A = [.......]; B = [.......]; C=30; nA = size(A,1); nB = size(B,1);

T = [A ; B];                        % training points
y = [ones(nA,1) ; -ones(nB,1)];     % labels
l = length(y);
Q = zeros(l,l);

for i = 1 : l
    for j = 1 : l
        Q(i,j) = y(i)*y(j)*(T(i,:))*T(j,:)';  % (minus) Dual Hessian
    end
end

% solve the problem
la = quadprog(Q,-ones(l,1),[ ],[ ],y',0,zeros(l,1),C*ones(l,1));

w = zeros(2,1);                   % compute vector w
for i = 1 : l
    w = w + la(i)*y(i)*T(i,:)';
end
```

```
indpos = find(la > 0.001) ;      % compute scalar b
ind = find(la(indpos) < C - 10^(-3));
  i = indpos(ind(1)) ;
 b = 1/y(i) - w'*T(i,:)'
```

We obtain the dual optimal solution:

```
la =

   30.0000
    0.0000
    0.0000
    0.0000
    0.0000
   30.0000
    0.0000
    0.0000
   30.0000
   11.2727
    0.0000
   30.0000
   30.0000
    0.0000
    0.0000
    0.0000
   26.3038
    0.0000
   14.9690
    0.0000



w =

   -0.8487
   -0.6539
   b =
     7.4975
```

The optimal hyperplane has equation $w^T x + b = -0.8487x_1 - 0.6539x_2 + 7.4975 = 0$.

(b) Consider the dual optimal solution λ^* and denote by (w^*, b^*, ξ^*) an optimal solution of (1). By the complementary slackness conditions,

$$\begin{cases} \lambda_i^* \left[1 - y^i((w^*)^T x^i + b^*) - \xi_i^*\right] = 0 \\ (30 - \lambda_i^*)\xi_i^* = 0 \end{cases} \tag{2}$$

it follows that a necessary condition for a point x^i to be misclassified is that $\lambda_i^* = 30$. We find that $\lambda_i^* = 30$, for $i = 1, 6, 9, 12, 13$, which correspond to the points

$$x^1 = (8, 1.5) \in A, \quad x^6 = (9.5, 0.34) \in A, \quad x^9 = (2.75, 6.75) \in A, \quad x^{12} = (5, 2) \in B, \quad x^{13} = (7.51, 2.55) \in B$$

The points x^1, x^6, x^{12} are misclassified, being $w^T x^i + b < 0,\ i = 1, 6,\ w^T x^{12} + b > 0$. Note that x^9, x^{13} are not misclassified being $w^T x^9 + b > 0$, $w^T x^{13} + b < 0$, in fact in this cases, we have the errors $\xi_9, \xi_{13}^* < 1$.

(c) The new point $\bar{x}^T = (4, 3)$ is labeled 1, since $w^T \bar{x} + b = 2.140 > 0$.

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

$$\begin{cases} \min\ (x_1 - x_2^2,\ x_2) \\ x_1 \geq 0 \\ x_2 \geq 0 \\ x \in \mathbb{R}^2 \end{cases}$$

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

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

(c) Does (P) admit a Pareto minimum point?

(d) Find a suitable subset of weak Pareto minima and, in case they exist, of Pareto minima.

SOLUTION

(a) The problem is not convex, since the objective function $f_1(x_1, x_2) = x_1 - x_2^2$ is not convex.

(b) Since the problem

$$\begin{cases} \min\ x_2 \\ x_1 \geq 0 \\ x_2 \geq 0 \\ x \in \mathbb{R}^2 \end{cases}$$

admits the points $\{(x_1, x_2) : x_2 = 0, x_1 \geq 0\}$ as global minima, all these points are weak minima for (P).

(c) Let X be the feasible set of (P). We notice that $(0, 0)$ is a Pareto minimum for (P), in fact, applying the definition, the systems

$$\begin{cases} f_1(0,0) - f_1(x_1, x_2) > 0 \\ f_2(0,0) - f_2(x_1, x_2) \geq 0 \\ x \in X \end{cases} \quad \text{and} \quad \begin{cases} f_1(0,0) - f_1(x_1, x_2) \geq 0 \\ f_2(0,0) - f_2(x_1, x_2) > 0 \\ x \in X \end{cases}$$

are both impossible. Namely, the two systems are:

$$\begin{cases} -x_1 + x_2^2 > 0 \\ -x_2 \geq 0 \\ x \in X \end{cases} \quad \text{and} \quad \begin{cases} -x_1 + x_2^2 \geq 0 \\ -x_2 > 0 \\ x \in X \end{cases}$$

and it is immediate to see that they are impossible.

(d) Let us compute the set $f(X) := \{(y_1, y_2) : y_1 = x_1 - x_2^2,\ y_2 = x_2, x_1 \geq 0,\ x_2 \geq 0\}$. Then

$$\begin{cases} y_1 = x_1 - y_2^2 \\ y_2 = x_2 \\ x_1 \geq 0,\ x_2 \geq 0 \end{cases}$$

which implies $f(X) := \{(y_1, y_2) : y_1 \geq -y_2^2,\ y_2 \geq 0\}$. It can be checked, graphically, that the optimality condition for a minimum of $f(X)$, namely,

$$f(X) \cap ((y_1, y_2) - \mathbb{R}^2_+)) = \{(y_1, y_2)\}$$

leads to the points

$$Min(f(X) = \{(y_1, y_2) : y_1 = -y_2^2,\ y_2 \geq 0\}$$

while the optimality condition for a weak minimum of $f(X)$, namely,

$$f(X) \cap ((y_1, y_2) - int\mathbb{R}^2_+) = \emptyset$$

leads to the points

$$WMin(f(X)) = Min(f(X)) \cup \{(y_1, y_2) : y_1 \geq 0,\ y_2 = 0\}$$

Therefore:

$$Min(P) = \{(x_1, x_2) : x_1 = 0,\ x_2 \geq 0\}$$

while

$$WMin(P)) = Min(P) \cup \{(x_1, x_2) : x_1 \geq 0,\ x_2 = 0\}$$

4) Consider the following matrix game:

$$C = \begin{pmatrix} -2 & 0 & 1 & 4 \\ 1 & 2 & -1 & 2 \\ 2 & 1 & 3 & 2 \end{pmatrix}$$

(a) Find the strictly dominated strategies, if any, and reduce the cost matrix 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.

(d) Is $\hat{x} = (0, \frac{3}{4}, \frac{1}{4})$, $\hat{y} = (0, 0, 0, 1)$ a mixed strategies Nash equilibrium? Justify the answer.

SOLUTION

(a) No strictly dominated strategy exists since no row is strictly (componentwise) greater than another and no columns is strictly (componentwise) minor than another.

(b) We observe that $c_{11}, c_{12}, c_{23}, c_{24}, c_{34}$, are the minima on the columns of the matrix C, while $c_{14}, c_{22}, c_{24}, c_{33}$ are the maxima on the rows. Therefore c_{24} corresponds to the pure strategies Nash equilibrium $(2, 4)$.

(c) The optimization problem associated with Player 1 is

$$\begin{cases} \min\ v \\ v \geq -2x_1 + x_2 + 2x_3 \\ v \geq 2x_2 + x_3 \\ v \geq x_1 - x_2 + 3x_3 \\ v \geq 4x_1 + 2x_2 + 2x_3 \\ x_1 + x_2 + x_3 = 1 \\ x \geq 0 \end{cases} \tag{3}$$

The previous problem can be solved by Matlab.

Matlab solution

```
C=[-2,0,1,4; 1 2 -1 2; 2 1 3 2 ]

m = size(C,1);
n = size(C,2);
c=[zeros(m,1);1];
A= [C', -ones(n,1)]; b=zeros(n,1); Aeq=[ones(1,m),0]; beq=1;
lb= [zeros(m,1);-inf]; ub=[ ];
[sol,Val,exitflag,output,lambda] = linprog(c, A,b, Aeq, beq, lb, ub);
x = sol(1:m)
y = lambda.ineqlin
```

We obtain the optimal solution $(\bar{x}, \bar{v}) = (0, \frac{1}{4}, \frac{3}{4}, 2)$. The optimal solution of the dual of (3) is given by $(\bar{y}, \bar{w}) = (0, 0, 0, 1, 2)$. y can be found in the vector *lambda.ineqlin* given by the Matlab function linprog.

Therefore,

$$(x_1, x_2, x_3) = (0, \frac{1}{4}, \frac{3}{4}), \quad (y_1, y_2, y_3, y_4) = (0, 0, 0, 1),$$

is a mixed strategies Nash equilibrium.

(d) We first note that $\hat{x}^T C \hat{y} = \bar{x}^T C \bar{y} = \bar{v} = 2$ and $\hat{x} \in X$. Since $(\hat{x}, \bar{v})$ is feasible for (P1), then $(\hat{x}, \bar{v})$ is optimal for (P1). Moreover, since $(\hat{y}, 2) = (\bar{y}, 2)$ is a dual solution of (P1), then $(\hat{x}, \hat{y})$ is a mixed strategies Nash equilibrium.