1) Consider the unconstrained optimization problem

$$\begin{cases} \min\ 3x_1^2 + x_2^2 + 2x_3^2 + 2x_4^2 + x_1x_2 + 2x_1x_4 + 2x_3x_4 + x_1 - x_2 + 2x_3 - 3x_4 \\ x \in \mathbb{R}^4 \end{cases}$$

(a) Apply the conjugate gradient method, with starting point $x^0 = (1, 0, 0, 0)$ and using $\|\nabla f(x)\| < 10^{-6}$ 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

$$Q = \begin{pmatrix} 6 & 1 & 0 & 2 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 4 & 2 \\ 2 & 0 & 2 & 4 \end{pmatrix} \qquad c^T = (1, -1, 2, -3)$$

Matlab solution

```
Q = [6 1 0 2;1 2 0 0;0 0 4 2;2 0 2 4]

c = [ 1 -1 2 -3]';


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

%% Parameters

x0 = [1,0,0,0]';
tolerance = 10^(-6);


%% Conjugate Gradient method

% starting point
x = x0;
X=[ ];
for ITER=1:10
    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
    if ITER == 1
        d = -g;
    else
        beta = (g'*Q*d_prev)/(d_prev'*Q*d_prev);
        d = -g + beta*d_prev;
    end

    %   step size
    t = (-g'*d)/(d'*Q*d);

    %   new point

    x = x + t*d;
```

```
    d_prev = d ;
end
 x
 v
 norm(g)
 ITER
```

We obtain the following solution:

```
Eigenvalues of Q:

ans =

    1.2984
    2.0000
    5.0000
    7.7016


x =

   -1.0000
    1.0000
   -1.5000
    2.0000



ITER =

    5

>> X(3:5,:)

    3.0000   -0.9814    0.5910   -1.0245    1.9476   -4.9333    2.1535
    4.0000   -0.9787    0.8671   -1.5425    2.0239   -5.4801    0.2817
    5.0000   -1.0000    1.0000   -1.5000    2.0000   -5.5000    0.0000
```

The iterations of the algorithm are 5 (the effective ones are 4). In particular, the gradient norm evaluated at the final point is: $3 \cdot 10^{-15}$.

(b) The found point $\bar{x} = (-1, 1, -1.5, 2)$ is 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 $\bar{x} = -Q^{-1}c$, which coincides with $\bar{x}$.

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

$$A = \begin{pmatrix} 9 & 1 \\ 6 & 2 \\ 7 & 0.5 \\ 2.5 & 0.4 \\ 0.9 & 8.5 \\ 9.5 & 0.3 \\ 4.3 & 3.8 \\ 2.8 & 4.8 \\ 2.75 & 6.75 \\ 6.5 & 2.5 \end{pmatrix}, \quad B = \begin{pmatrix} 11 & 2.5 \\ 5 & 2 \\ 7.5 & 2.5 \\ 5 & 7 \\ 9 & 9.6 \\ 7.4 & 3.5 \\ 8.3 & 2.4 \\ 9.3 & 3.4 \\ 6.4 & 5 \\ 3.5 & 8.5 \end{pmatrix}$$

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

(b) Solve the dual problem with parameter $C = 10$ 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 $(6, 5)$.

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

$$\begin{cases} \min\limits_{w,b,\xi} \ \dfrac{1}{2}\|w\|^2 + 10 \sum\limits_{i=1}^{20} \xi_i \\ 1 - y^i(w^T x^i + b) \le \xi_i, \qquad \forall\, i = 1, \dots, 20 \\ \xi_i \ge 0, \qquad \forall\, i = 1, \dots, 20 \end{cases} \tag{1}$$

(b) The dual problem of (1) is

$$\begin{cases} \max\limits_{\lambda} \ -\dfrac{1}{2} \sum\limits_{i=1}^{20} \sum\limits_{j=1}^{20} y^i y^j (x^i)^T x^j \lambda_i \lambda_j + \sum\limits_{i=1}^{20} \lambda_i \\ \sum\limits_{i=1}^{20} \lambda_i y^i = 0 \\ 0 \le \lambda_i \le 10 \qquad i = 1, \dots, 20 \end{cases}$$

Matlab solution

```
A = [.......]; B = [.......]; C=10; 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 =

   10.0000
    0.0000
    0.0000
    0.0000
    0.0000
   10.0000
    0.0000
    0.0000
    9.2605
   10.0000
    0.0000
   10.0000
   10.0000
    0.0000
    0.0000
    9.2605
   10.0000
    0.0000
    0.0000
    0.0000
```

```
w =

   -1.0614
   -0.9033

    b =

     10.0162
```

The optimal hyperplane has equation $w^T x + b = -1.0614x_1 - 0.9033x_2 + 10.0162 = 0$.

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

$$\begin{cases} \lambda_i^* \left[1 - y^i((w^*)^T x^i + b^*) - \xi_i^*\right] = 0 \\ (10 - \lambda_i^*)\xi_i^* = 0 \end{cases} \tag{2}$$

it follows that a necessary condition for a point x^i to be misclassified is that $\lambda_i^* = 10$. We find that $\lambda_i^* = 10$, for $i = 1, 6, 10, 12, 13, 17$, which correspond to the points

$$x^1 = (9, 1.5) \in A, \quad x^6 = (9.5, 0.3) \in A, \quad x^{10} = (6.5, 2.5) \in A, \quad x^{12} = (5, 2) \in B, \quad x^{13} = (7.5, 2.5) \in B, \quad x^{17} = (8.3, 2.4) \in B$$

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^{10}, x^{13}, x^{17} are not misclassified being $w^T x^{10} + b > 0$, $w^T x^{13} + b < 0$, $w^T x^{17} + b < 0$, in fact in this cases, we have the errors $\xi^*_{10}, \xi^*_{13}, \xi^*_{17} < 1$.

(c) The new point $\bar{x}^T = (6, 5)$ is labeled -1, since $w^T \bar{x} + b = -0.8689 < 0$.

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

$$\begin{cases} \min\ (x_1^2 - 3x_2,\ x_2^2 + x_1) \\ x_1 - x_2 \leq 0 \\ -x_1 \leq 0 \\ x \in \mathbb{R}^2 \end{cases}$$

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

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

(c) Find a suitable subset of weak Pareto minima.

(d) Find a suitable subset of Pareto minima.

SOLUTION

(a) The problem is convex, since the objective functions are convex and the constraints are linear. Therefore the set of weak minima coincides with the set of solutions of the scalarized problems (P_{α_1}), where $0 \leq \alpha_1 \leq 1$, i.e.

$$\begin{cases} \min\ \alpha_1 x_1^2 + (1-\alpha_1)x_2^2 + (1-\alpha_1)x_1 - 3\alpha_1 x_2 =: \psi_{\alpha_1}(x) \\ x_1 - x_2 \leq 0 \\ -x_1 \leq 0 \end{cases}$$

Moreover, for every $0 < \alpha_1 < 1$ we obtain a Pareto minimum.

(b) Since the problem

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

admits the point $(x_1, x_2) = (0,0)$ as unique global minimum, $(0,0)$ is a Pareto minimum for (P).

(c) -(d)

Let us consider the KKT conditions for (P_{α_1}) which is convex,differentiable and fulfils the Abadie constraints qualifications. Therefore, the following KKT system provides a necessary and sufficient condition for an optimal solution of (P_{α_1}):

$$\begin{cases} 2\alpha_1 x_1 + 1 - \alpha_1 + \lambda_1 - \lambda_2 = 0 \\ 2(1-\alpha_1)x_2 - 3\alpha_1 - \lambda_1 = 0 \\ \lambda_1(x_1 - x_2) = 0 \\ \lambda_2(-x_1) = 0 \\ x_1 - x_2 \leq 0, -x_1 \leq 0 \\ \lambda_i \geq 0,\ i = 1,2 \\ 0 \leq \alpha_1 \leq 1, \end{cases} \tag{3}$$

In particular, notice that for $\alpha_1 = 1$ the previous system is impossible, by the second equation.
For $\alpha_1 = 0$, the system becomes

$$\begin{cases} 1 + \lambda_1 - \lambda_2 = 0 \\ 2x_2 - \lambda_1 = 0 \\ \lambda_1(x_1 - x_2) = 0 \\ \lambda_2(-x_1) = 0 \\ x_1 - x_2 \leq 0, -x_1 \leq 0 \\ \lambda_i \geq 0,\ i = 1,2 \end{cases} \tag{4}$$

Note that $\lambda_2 = 1 + \lambda_1 > 0$ implies $x_1 = 0$ so that the system (4) admits the unique solution $\bar{x} = (0,0)$, $\bar{\lambda} = (0,1)$ so that $\bar{x}$ is a Pareto minimum for (P).

Consider the case $0 < \alpha_1 < 1$. then system (3) becomes:

$$\begin{cases} x_1 = \dfrac{-1+\alpha_1-\lambda_1+\lambda_2}{2\alpha_1} \\ x_2 = \dfrac{3\alpha_1+\lambda_1}{2(1-\alpha_1)} \\ \lambda_1(x_1-x_2)=0 \\ \lambda_2(-x_1)=0 \\ x_1-x_2 \leq 0, -x_1 \leq 0 \\ \lambda_i \geq 0,\ i=1,2 \\ 0<\alpha_1<1, \end{cases} \tag{5}$$

Observe that λ_1 and λ_2 cannot be simultaneously striclty positive nor simultameously zero. In fact in the first case we would have $x=(0,0)$ which contradicts the second equation which implies $x_2>0$. In the second case, by the first equation we have $x_1 = \frac{\alpha_1-1}{2\alpha_1} < 0$, which is impossible.

Now assume $\lambda_1 = 0$, $\lambda_2 \neq 0$. Then the system becomes

$$\begin{cases} \lambda_2 = 1-\alpha_1 \\ x_2 = \dfrac{3\alpha_1}{2(1-\alpha_1)} \\ x_1 = 0 \\ -x_2 \leq 0, \\ \lambda_i \geq 0,\ i=1,2 \\ 0<\alpha_1<1, \end{cases} \tag{6}$$

Notice that in the case $\lambda_1 \neq 0$, $\lambda_2 = 0$, the system (5) is impossible.

Therefore:

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

and

$$WMin(P) = Min(P)$$

4) Consider the following matrix game:

$$C = \begin{pmatrix} 0 & -1 & 2 & 4 \\ 3 & 2 & -1 & 5 \\ -2 & 4 & 3 & 1 \end{pmatrix}$$

(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 the found mixed strategies Nash equilibrium unique? Justify the answer.

SOLUTION

(a) Strategy 1 of Player 2 is dominated by strategy 4.

(b) We observe that $c_{31}, c_{12}, c_{23}, c_{24}$, are the minima on the columns of the matrix C, while $c_{14}, c_{22}, c_{24}, c_{32}$ are the maxima on the rows. 5Therefore c_{24} corresponds to the pure strategies Nash equilibrium $(2, 4)$. Since there are no common couples, no pure strategies Nash equilibrium exists.

(c) The optimization problem associated with Player 1 is

$$\begin{cases} \min\ v \\ v \geq 3x_2 - 2x_3 \\ v \geq -x_1 + 2x_2 + 4x_3 \\ v \geq 2x_1 - x_2 + 3x_3 \\ v \geq 4x_1 + 5x_2 + x_3 \\ x_1 + x_2 + x_3 = 1 \\ x \geq 0 \end{cases} \tag{7}$$

The previous problem can be solved by Matlab.

Matlab solution

```
C=[0,-1,2,4; 3 2 -1 5; -2 4 5 1 ]

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 $(\bar{x}, \bar{v}) = (\frac{3}{10}, \frac{1}{10}, \frac{3}{5}, 2.3)$. The optimal solution of the dual of (7) is given by $(\bar{y}, \bar{w}) = (0, \frac{1}{5}, \frac{7}{20}, \frac{9}{20}, 2.3)$. y can be found in the vector *lambda.ineqlin* given by the Matlab function linprog.
Therefore,

$$(x_1, x_2, x_3) = (\frac{3}{10}, \frac{1}{10}, \frac{3}{5}, 2.3), \quad (y_1, y_2, y_3, y_4) = (0, \frac{1}{5}, \frac{7}{20}, \frac{9}{20}),$$

is a mixed strategies Nash equilibrium.

(d) By exploiting the complementarity conditions for problem (7), all the optimal solutions of (7) solve the system

$$\begin{cases} v \geq 3x_2 - 2x_3 \\ v = -x_1 + 2x_2 + 4x_3 \\ v = 2x_1 - x_2 + 3x_3 \\ v = 4x_1 + 5x_2 + x_3 \\ x_1 + x_2 + x_3 = 1 \\ v = 2.3 \\ x \geq 0 \end{cases} \tag{8}$$

that can be shown to have the unique solution $(\bar{x}, \bar{v})$.
Consider the problem related to Player 2, namely

$$\begin{cases} \max \ w \\ w \leq -y_2 + 2y_3 + 4y_4 \\ w \leq 3y_1 + 2y_2 - y_3 + 5y_4 \\ w \leq -2y_1 + 4y_2 + 3y_3 + y_4 \\ y_1 + y_2 + y_3 + y_4 = 1 \\ y \geq 0 \end{cases} \tag{9}$$

By the complementarity relations for problem (9), all the optimal solutions of (9) solve the system

$$\begin{cases} w = -y_2 + 2y_3 + 4y_4 \\ w = 3y_1 + 2y_2 - y_3 + 5y_4 \\ w = -2y_1 + 4y_2 + 3y_3 + y_4 \\ y_1 + y_2 + y_3 + y_4 = 1 \\ w = 2.3 \\ y \geq 0 \end{cases} \tag{10}$$

that can be shown to have the unique solution $(\bar{y}, \bar{w} = 2.3)$. This shows that $(\bar{x}, \bar{y})$ is the unique Nash equilibrium.