1) Consider the unconstrained optimization problem

$$\begin{cases} \min\ 3x_1^2 + x_2^2 + 2x_3^2 + 4x_4^2 + x_1x_2 + 2x_1x_4 + 2x_3x_4 - x_1 + 8x_2 + 6x_3 + 9x_4 \\ x \in \mathbb{R}^4 \end{cases}$$

(a) Apply the conjugate gradient method with starting point $x^0 = (0, 0, 0, 0)$ and using $\|\nabla f(x)\| < 10^{-6}$ as stopping criterion. How many iterations are needed by the algorithm? Write the vector found at each iteration.

(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., $f(x) = (1/2)x^T Qx + c^T x$ with

$$Q = \begin{pmatrix} 6 & 1 & 0 & 2 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 4 & 2 \\ 2 & 0 & 2 & 8 \end{pmatrix} \qquad c^T = (-1, 8, 6, 9)$$

Matlab solution

```
Q = [6 1 0 2;1 2 0 0;0 0 4 2;2 0 2 8];
c = [-1 8 6 9]';

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

x0 = [0,0,0,0]';         % Starting point
tolerance = 10^(-6);     % Parameters

x = x0;
X = [];
for ITER = 1:10
    v = 0.5*x'*Q*x + c'*x;
    g = Q*x + c ;

    X=[X;ITER,x',v];
    if norm(g) < tolerance   % stopping criterion
       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

    t = (-g'*d)/(d'*Q*d);    % step size
    x = x + t*d;             % new point
    d_prev = d;
end
 X
 disp('Gradient norm:')
 norm(g)
```

We obtain the following solution:

```
Eigenvalues of Q:
ans =
    1.7070
    3.0000
    5.5077
    9.7853
```

```
X =
    1.0000         0         0         0         0         0
    2.0000    0.1670   -1.3358   -1.0018   -1.5028  -15.1945
    3.0000    1.6166   -3.9878   -1.4191   -1.1575  -26.2254
    4.0000    1.2719   -4.6713   -1.0360   -1.1446  -27.5798
    5.0000    1.3623   -4.6812   -0.8768   -1.2464  -27.6449
```

In particular, the gradient norm evaluated at the final point is:

```
ans =

   4.4409e-16
```

The effective iterations of the algorithm are 4, since in the first one we have considered the initial point x^0.

(b) The found point $x = (1.3623 - 4.6812 - 0.8768 - 1.2464)$ is a global minimum since the objective function is strongly convex: in fact the eigenvalues of the Hessian of f are all strictly positive.

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

$$A = \begin{pmatrix} 6.55 & 0.85 \\ 6.55 & 1.71 \\ 7.06 & 0.31 \\ 2.76 & 0.46 \\ 0.97 & 8.23 \\ 9.5 & 0.34 \\ 4.38 & 3.81 \\ 1.86 & 4.89 \\ 2.76 & 6.79 \\ 6.55 & 1.22 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 9.59 & 3.40 \\ 5.85 & 2.23 \\ 7.51 & 2.55 \\ 5.05 & 7 \\ 8.9 & 9.59 \\ 8.40 & 2.54 \\ 8.14 & 2.43 \\ 9.3 & 3.45 \\ 6.16 & 4.73 \\ 3.51 & 8.30 \end{pmatrix}$$

(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 $(5, 5)$.

SOLUTION

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

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

Given $C > 0$, the formulation of the linear SVM with soft margin is

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

(b) The dual problem of (1) is

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

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 =

    0.0000
    6.5691
    0.0000
    0.0000
    0.0000
   10.0000
    0.0000
    0.0000
    7.8340
    0.0000
    0.0000
   10.0000
   10.0000
    0.0000
    0.0000
    0.0000
    0.0000
    0.0000
    4.4031
    0.0000


w =
   -1.0735
   -0.8009
b =
    9.4008
```

The optimal hyperplane has equation $w^T x + b = -1.0735x_1 - 0.8009x_2 + 9.4008 = 0$.

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

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

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

$$x^6 = (9.5, 0.34) \in A, \quad x^{12} = (5.85, 2.23) \in B, \quad x^{13} = (7.51, 2.55) \in B$$

The first two points are misclassified, being $w^T x^6 + b < 0$, $w^T x^{12} + b > 0$. Note that x^{13} is not misclassified being $w^T x^{13} + b < 0$, in fact in this case, even though $\lambda_{13}^* > 0$, we have that the error $\xi_{13}^* < 1$.

(c) The new point $\bar{x}^T = (5, 5)$ is labeled 1, since $w^T \bar{x} + b = 0.029 > 0$.

3) Consider the following multiobjective optimization problem:

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

(a) Prove that the problem admits a Pareto minimum point.

(b) Find the set of all weak Pareto minima.

(c) Find a suitable subset of Pareto minima.

SOLUTION

(a) We preliminarly observe that the problem is convex, since the objective and the constraint functions are convex. Therefore the set of weak minima coincides with the set of solutions of the scalarized problems (P_{α_1}), where $0 \le \alpha_1 \le 1$, i.e.

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

with $0 \le \alpha_1 \le 1$ and for every $0 < \alpha_1 < 1$ we obtain a Pareto minimum. Note that in this last case the objective function $\psi_{\alpha_1}(x)$ of P_{α_1} is strongly convex so that P_{α_1} admits a unique global solution for every $0 < \alpha_1 < 1$, which is a Pareto minimum.

(b) - (c) We note that for $\alpha_1 = 0$, the problem P_{α_1} is unbounded, indeed, at any feasible point $(x_1, 0)$, $d = (1, 0)$ is a descent direction which is also a recession direction for the polyhedron K.

P_{α_1} can be solved by Matlab for $0 < \alpha_1 \le 1$:

```
Q1 = [2 0; 0 0];
Q2 = [0 0; 0 2] ;
c1=[0 0]';
c2=[-1 0]';

A =[ -1 -1];
b = -1;

% solve the scalarized problem with  alfa1 in (0,1]

MINIMA=[ ]; LAMBDA=[ ]; % First column: value of alfa1

for alfa1 = 0.01 : 0.001 : 1
[x,fval,exitflag,output,lambda] = quadprog(alfa1*Q1+(1-alfa1)*Q2,alfa1*c1+(1-alfa1)*c2,A,b) ;
MINIMA=[MINIMA; alfa1 x'];
LAMBDA=[LAMBDA;alfa1,lambda.ineqlin'];
end

  plot(MINIMA(:,2),MINIMA(:,3), 'r*')
```

- For $0 < \alpha_1 \le 1/3$, we obtain the set of points $\{(x_1, 0), x_1 \ge 1\}$ which are Pareto minima.
- For $1/3 < \alpha_1 < 1$, we obtain the set of points $\{(x_1, x_2), x_1 + x_2 = 1, 0 < x_1 < 1\}$ which are Pareto minima.
- For $\alpha_1 = 1$, we obtain the set of points $\{(0, x_2), x_2 \ge 1\}$ which are Weak Pareto minima.

We note that the previous solutions can also be obtained by solving the KKT conditions for (P_{α_1}) which is convex,differentiable and fulfils the Abadie constraints qualifications. Therefore, the following system provides a necessary and sufficient condition for an optimal solution of (P_{α_1}):

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

In particular, notice that for $\alpha_1 = 0$ the previous system is impossible, which agrees with the fact that (P_{α_1}) has no solutions for $\alpha_1 = 0$, as already observed.

For $\alpha_1 = 1$, the system becomes:

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

and we obtain the set of points $\{(0, x_2), x_2 \geq 1\}$ which are Weak Pareto minima.

For $0 < \alpha_1 < 1$, the system becomes:

$$\begin{cases} x_1 = \frac{1-\alpha_1+\lambda}{2\alpha_1} \\ x_2 = \frac{\lambda}{2(1-\alpha_1)} \\ \lambda(-x_1 - x_2 + 1) = 0 \\ -x_1 - x_2 + 1 \leq 0, \lambda \geq 0 \\ 0 < \alpha_1 < 1, \end{cases} \tag{3}$$

For solving (3) we distinguish the cases $I)$ $\lambda = 0$ and $II)$ $\lambda > 0$. In case $I)$, (3) becomes:

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

which leads to the set of minimum points: $\{(x_1, 0), x_1 \geq 1\}$, with $0 < \alpha_1 \leq 1/3$.

In case $II)$,(3) becomes:

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

with solutions:

$$\begin{cases} x_1 = \frac{3(1-\alpha_1)}{2} \\ x_2 = \frac{3\alpha_1-1}{2} \\ \lambda = (3\alpha_1 - 1)(1 - \alpha_1) \\ \lambda > 0 \\ 0 < \alpha_1 < 1, \end{cases}$$

which leads to the set of minimum points: $\{(x_1, x_2), x_1 + x_2 = 1, 0 < x_1 < 1\}$, with $1/3 < \alpha_1 < 1$.

4) Consider the following bimatrix game:

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

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

(b) Find a mixed strategies Nash equilibrium.

SOLUTION

(a) Strategy 1 of Player 2 is dominated by Strategy 3, so that column 1 in the two matrices can be deleted. The reduced game is given by the matrices

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

Now, it is simple to show that (1,2) and (2,3) are pure strategies Nash equilibria. Indeed, the minima on the columns of C_1^R, (i.e., 0 and -1), are obtained in correspondence of the minima on the rows of C_2^R, (i.e., 1 and 1) and are related to the components $(1,2)$ $(2,3)$ of the given matrices C_1 and C_2.

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

$$\left\{ \begin{array}{l} \min\ x^T C_1^R y = 3x_2y_2 + (2x_1 - x_2)y_3 \\ x_1 + x_2 = 1 \\ x_1, x_2 \geq 0 \end{array} \right. \qquad \equiv \left\{ \begin{array}{l} \min\ (6y_2 - 3)x_2 - 2y_2 + 2 \\ 0 \leq x_1 \leq 1 \end{array} \right. \qquad (P_1(y_2))$$

Then, the best response mapping associated with $P_1(y_2)$ is:

$$B_1(y_2) = \left\{ \begin{array}{cl} 0 & \text{if } y_2 \in (1/2, 1] \\ [0,1] & \text{if } y_2 = 1/2 \\ 1 & \text{if } y_2 \in [0, 1/2) \end{array} \right.$$

Similarly, the optimization problem associated with Player 2 is

$$\left\{ \begin{array}{l} \min\ x^T C_2^R y = (x_1 + 4x_2)y_2 + (2x_1 + x_2)y_3 \\ y_2 + y_3 = 1 \\ y_2, y_3 \geq 0 \end{array} \right. \qquad \equiv \left\{ \begin{array}{l} \min\ (4x_2 - 1)y_2 + 2 - x_2 \\ 0 \leq y_2 \leq 1 \end{array} \right. \qquad (P_2(x_2))$$

Then, the best response mapping associated with $P_2(x_2)$ is:

$$B_2(x_2) = \left\{ \begin{array}{cl} 0 & \text{if } x_2 \in (1/4, 1] \\ [0,1] & \text{if } x_1 = 1/4 \\ 1 & \text{if } x_1 \in [0, 1/4) \end{array} \right.$$

The couples (x_2, y_2) such that $x_2 \in B_1(y_2)$ and $y_2 \in B_2(x_2)$ are

1. $x_2 = 0, y_2 = 1$,
2. $x_2 = \frac{1}{4}, y_2 = \frac{1}{2}$,
3. $x_2 = 1, y_2 = 0$,

so that, recalling that $y_1 = 0$,

- $(x_1, x_2) = (0, 1)$, $\quad (y_1, y_2, y_3) = (0, 0, 1)$, is a pure strategies Nash equilibrium,
- $(x_1, x_2) = (\frac{3}{4}, \frac{1}{4})$, $\quad (y_1, y_2, y_3) = (0, \frac{1}{2}, \frac{1}{2})$, is a mixed strategies Nash equilibrium,
- $(x_1, x_2) = (1, 0)$, $\quad (y_1, y_2, y_3) = (0, 1, 0)$, is a pure strategies Nash equilibrium.