1) Consider the unconstrained optimization problem

$$\begin{cases} \min \ 2x_1^2 + x_2^2 - x_1x_2 + e^{x_1+2x_2} \\ x \in \mathbb{R}^2 \end{cases}$$

(a) Prove that the problem admits a global minimum;

(b) Apply the gradient method with an inexact line search, setting $\bar{t} = 1$,$\alpha = 0.1$, $\gamma = 0.8$, with starting point $x^0 = (-10, 8)$ and using $\|\nabla f(x)\| < 10^{-6}$ as stopping criterion. How many iterations are needed by the algorithm? Write explicitly the vectors found at the last three iterations.

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

SOLUTION

(a) The objective function $f(x) = f_1(x) + f_2(x)$, where $f_1 = 2x_1^2 + x_2^2 - x_1x_2$ is strongly convex and $f_2(x) = e^{x_1+2x_2}$ is convex, being $f_2 = \psi \circ g$, with $\psi(y) = e^y$ convex and $g(x) = x_1 + 2x_2$ linear. Therefore $f_1 + f_2$ is strongly convex being the sum of a strongly convex plus a convex function. Consequently, f admits a unique global minimum point.

(b) We notice that

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4x_1 - x_2 + e^{x_1+2x_2} \\ -x_1 + 2x_2 + 2e^{x_1+2x_2} \end{pmatrix}$$

Matlab solution

```
%% Data

alpha = 0.1;
gamma = 0.8;
tbar = 1;
x0 = [-10;8];
tolerance = 10^(-6) ;

X=[ ];

ITER = 0 ;
x = x0;

while true
    [v, g] = f(x);

    X=[X;ITER,x(1),x(2),v,norm(g)];

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

    d = -g;                         % search direction


     t = tbar ;                       % Armijo inexact line search
    while f(x+t*d) > v + alpha*g'*d*t
        t = gamma*t ;
    end


    x = x + t*d ;                   % new point
    ITER = ITER + 1 ;
    end

disp('optimal solution')
x
v
norm(g)
ITER
```

```
function [v, g] = f(x)

v = 2*x(1)^2 + x(2)^2 - x(1)*x(2) + exp(x(1)+2*x(2)) ;

g = [4*x(1)-x(2)+exp(x(1)+2*x(2));
      -x(1)+2*x(2)+2*exp(x(1)+2*x(2))];

end
```

We obtain the following solution:

```
x =

   -0.1952
   -0.4393


v =

    0.5251


ans =

   6.6986e-07


ITER =

    24
```

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

```
ans =

   6.6986e-07
```

The iterations of the algorithm are 24.

The vectors found at the last three iterations are:

```
-0.1952   -0.4393
-0.1953   -0.4393
-0.1952   -0.4393
```

(b) The found point $x = (-0.1952, -0.4393)$ is a global minimum since the objective function is strongly convex as shown in point (a).

2) Consider a regression problem with the following data set where the points $(x_i, y_i), i = 1, 31$, are given by the row vectors of the matrices:

$$\begin{pmatrix} -3.0000 & 4.58 \\ -2.8000 & 7.19 \\ -2.6000 & 8.22 \\ -2.4000 & 16.06 \\ -2.2000 & 16.42 \\ -2.0000 & 17.53 \\ -1.8000 & 11.48 \\ -1.6000 & 14.10 \\ -1.4000 & 16.82 \\ -1.2000 & 16.15 \\ -1.0000 & 11.68 \\ -0.8000 & 6.00 \\ -0.6000 & 7.82 \\ -0.4000 & 2.82 \\ -0.2000 & 2.71 \\ 0 & 1.16 \end{pmatrix} \qquad \begin{pmatrix} 0.2000 & -1.42 \\ 0.4000 & -3.84 \\ 0.6000 & -4.71 \\ 0.8000 & -8.15 \\ 1.0000 & -7.33 \\ 1.2000 & -13.64 \\ 1.4000 & -15.26 \\ 1.6000 & -14.87 \\ 1.8000 & -9.92 \\ 2.0000 & -10.50 \\ 2.2000 & -7.72 \\ 2.4000 & -11.78 \\ 2.6000 & -10.26 \\ 2.8000 & -7.13 \\ 3.0000 & -2.11 \end{pmatrix}$$

(a) Write the dual formulation of a nonlinear ε-SV regression model with $C = 5$, $\varepsilon = 3$ and a polynomial kernel $k(x, y) := (x^T y + 1)^4$;

(b) Solve the problem in (a) and find the regression function;

(c) Find the support vectors;

(d) Find the points of the data set that are outside the ε-tube, by making use of the dual solution.

SOLUTION

(a) Let $\ell = 31$, $(x_i, y_i))$, $i = 1, .., \ell$ be the i-th element of the data set, $C = 5$, $\varepsilon = 3$, $k(x, y) := (x^T y + 1)^4$. The dual formulation of a nonlinear ε-SV regression model is

$$\begin{cases} \max\limits_{(\lambda^+,\lambda^-)} & -\dfrac{1}{2}\sum\limits_{i=1}^{\ell}\sum\limits_{j=1}^{\ell}(\lambda_i^+ - \lambda_i^-)(\lambda_j^+ - \lambda_j^-)k(x_i, x_j) \\ & -\varepsilon\sum\limits_{i=1}^{\ell}(\lambda_i^+ + \lambda_i^-) + \sum\limits_{i=1}^{\ell} y_i(\lambda_i^+ - \lambda_i^-) \\ \sum\limits_{i=1}^{\ell}(\lambda_i^+ - \lambda_i^-) = 0 \\ \lambda_i^+, \lambda_i^- \in [0, C], i = 1, .., \ell \end{cases} \tag{1}$$

(b) **Matlab solution**

```
data = [
   -3.0000    4.58
   -2.8000     7.19
   -2.6000     8.22
   -2.4000    16.06
   -2.2000    16.42
   -2.0000    17.53
   -1.8000    11.48
   -1.6000    14.10
   -1.4000    16.82
   -1.2000    16.15
   -1.0000    11.68
   -0.8000     6.00
   -0.6000     7.82
   -0.4000     2.82
   -0.2000     2.71
         0     1.16
    0.2000    -1.42
    0.4000    -3.84
    0.6000    -4.71
    0.8000    -8.15
    1.0000    -7.33
    1.2000   -13.64
    1.4000   -15.26
    1.6000   -14.87
```

```
    1.8000   -9.92
    2.0000  -10.50
    2.2000   -7.72
    2.4000  -11.78
    2.6000  -10.26
    2.8000   -7.13
    3.0000   -2.11
        ];

x = data(:,1) ;
y = data(:,2) ;
l = length(x) ;

epsilon = 3 ;
C = 5;

X = zeros(l,l);
for i = 1 : l
    for j = 1 : l
        X(i,j) = kernel(x(i),x(j)) ;
    end
end
Q = [ X -X ; -X X ];
c = epsilon*ones(2*l,1) + [-y;y];

sol = quadprog(Q,c,[],[],[ones(1,l) -ones(1,l)],0,zeros(2*l,1),C*ones(2*l,1));
lap = sol(1:l);
lam = sol(l+1:2*l);

% compute b
ind = find(lap > 1e-3 & lap < C-1e-3);
if isempty(ind)==0
    i = ind(1);
    b = y(i) - epsilon;
    for j = 1 : l
        b = b - (lap(j)-lam(j))*kernel(x(i),x(j));
    end
else
    ind = find(lam > 1e-3 & lam < C-1e-3);
    i = ind(1);
    b = y(i) + epsilon ;
    for j = 1 : l
        b = b - (lap(j)-lam(j))*kernel(x(i),x(j));
    end
end


z = zeros(l,1);                                                % find regression and epsilon-tube
for i = 1 : l
    z(i) = b ;
    for j = 1 : l
        z(i) = z(i) + (lap(j)-lam(j))*kernel(x(i),x(j));
    end
end
zp = z + epsilon ;
zm = z - epsilon ;

sv = [find(lap > 1e-3);find(lam > 1e-3)];                      % find support vectors
sv = sort(sv);

plot(x,y,'b.',x(sv),y(sv),'ro',x,z,'k-',x,zp,'r-',x,zm,'r-');     % plot the solution

    disp('Support vectors')
```

```
[sv,x(sv),y(sv),lam(sv),lap(sv)]    % Indexes of support vectors, support vectors, lambda_-,lambda_+

function v = kernel(x,y)
p = 4 ;
v = (x'*y + 1)^p;
end
```

Let λ_- and λ_+ be the vectors given by the Matlab solutions lam, lap. In particular we find, $b = -0.008$.

The regression function is:

$$f(x) = \sum_{i=1}^{\ell}(\lambda_i^+ - \lambda_i^-)k(x_i, x) + b = \sum_{i=1}^{31}(\lambda_i^+ - \lambda_i^-)(x_i x + 1)^4 - 0.008$$

(c) We obtain the support vectors (columns 2-3) and corresponding λ_- and λ_+ (columns 4-5) :

```
ans =

    4.0000   -2.4000   16.0600    0.0000    1.2569
    7.0000   -1.8000   11.4800    5.0000    0.0000
    9.0000   -1.4000   16.8200    0.0000    1.3176
   10.0000   -1.2000   16.1500    0.0000    5.0000
   12.0000   -0.8000    6.0000    0.5720    0.0000
   23.0000    1.4000  -15.2600    4.7410    0.0000
   27.0000    2.2000   -7.7200    0.0000    5.0000
   29.0000    2.6000  -10.2600    2.2616    0.0000
```

(d) Consider the feasibility condition of the primal formulation of the regression problem:

$$y_i - f(x_i) - \varepsilon - \xi_i^+ \leq 0, \quad y_i - f(x_i) + \varepsilon + \xi_i^- \geq 0, \quad i = 1, .., \ell$$

If a point (x_i, y_i) is outside the ε-tube then $\xi_i^+ > 0$ or $\xi_i^- > 0$.

Given the dual optimal solution (λ_+, λ_-) of (1), we can find the errors ξ_i^+ and ξ_i^- associated with the point (x_i, y_i) by the complementarity conditions:

$$\begin{cases} \lambda_i^+ \left[y_i - f(x_i) - \varepsilon - \xi_i^+\right] = 0, & i = 1, .., \ell \\ \lambda_i^- \left[y_i - f(x_i) + \varepsilon + \xi_i^-\right] = 0, & i = 1, .., \ell \\ \xi_i^+(C - \lambda_i^+) = 0, & i = 1, .., \ell \\ \xi_i^-(C - \lambda_i^-) = 0, & i = 1, .., \ell \end{cases} \tag{2}$$

it follows that a necessary condition for a point (x_i, y_i) to be outside the ε-tube is that $\lambda_i^+ = C = 5$ or $\lambda_i^- = C = 5$. We find that $\lambda_i^- = 5$, for $i = 7$, $\lambda_i^+ = 5$, for $i = 10, 27$ which correspond to the points

$$(x_7, y_7) = (-1.8, 11.48), \quad (x_{10}, y_{10}) = (-1.2, 16.15), \quad (x_{27}, y_{27}) = (2.2, -7.72)$$

3) Consider the following unconstrained multiobjective optimization problem:

$$\begin{cases} \min \ f(x_1, x_2) = (x_1x_2, \ x_1^2 + 2x_2^2 - 2x_2) \\ x \in \mathbb{R}^2 \end{cases}$$

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

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

SOLUTION

(a) We observe that the function f_2 is strongly convex, so that its minimum point is a Pareto minimum of the given problem.

(b) Consider the scalarized problem (P_{α_1}), where $0 \le \alpha_1 \le 1$, i.e.

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

We note that the objective function $\psi_{\alpha_1}(x)$ is quadratic with Hessian given by

$$Q = \begin{pmatrix} 2(1-\alpha_1) & \alpha_1 \\ \alpha_1 & 4(1-\alpha_1) \end{pmatrix}$$

We observe that Q is positive definite for $0 \le \alpha_1 < \frac{8-2\sqrt{2}}{7} \approx 0.7388$ so that for such values P_{α_1} is strongly convex and admits a unique global solution which is a Pareto minimum.

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

```
Q1 = [0 1; 1 0];
Q2 = [2 0; 0 4] ;
c1=[0 0]’;
c2=[0 -2]’;


MINIMA=[ ]; SOL=[ ];

for alfa1 = 0 : 0.001 : 1
eigQalfa1= eig(alfa1*Q1+(1-alfa1)*Q2);
eigQ=[eigQ;alfa1,eigQalfa1’]; % Eigenvalues of the hessian of \psi_alfa1(x)

if (eigQ > 0.001)
[x,fval,exitflag,exitflag] = fminunc(@(x) 0.5*x’*(alfa1*Q1+(1-alfa1)*Q2)*x  +(alfa1*c1+(1-alfa1)*c2)’*x, [0,0]’

MINIMA=[MINIMA; alfa1 x’];

else

[x,fval,exitflag,exitflag] = fminunc(@(x) 0.5*x’(alfa1*Q1+(1-alfa1)*Q2)*x  +(alfa1*c1+(1-alfa1)*c2)’*x, [0,0]’)
SOL=[SOL; alfa1 x’];
end
end

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

For every $0 \le \alpha_1 \le 0.738$, the problem admits a unique solution which is a Pareto minimum (MINIMA(:,2:3)). In particular,

- For $\alpha_1 = 0$ we obtain the point $(0, 0.5)$
- For $\alpha_1 = 0.736$ we obtain the point $(-24.4839, 17.5645)$;
- For $\alpha_1 = 0.737$ we obtain $(-38.0695, 27.1704)$
- For $\alpha_1 = 0.738$ we obtain $(-85.7834, 60.9085)$
- For $\alpha_1 \ge 0.739$ the problem P_{α_1} is unbounded.

The previous values lead us to conjecture that the sequence of minima diverges as $\alpha_1 \to \frac{8-2\sqrt{2}}{7} \approx 0.7388$

We note that the previous solutions can also be obtained by the stationarity conditions for (P_{α_1}) which are necessary for a weak minimum point. Then, the following system provides a necessary condition for an optimal solution of (P_{α_1}) and such condition is also sufficient for $0 \leq \alpha_1 < \frac{8-2\sqrt{2}}{7}$:

$$\begin{cases} \alpha_1 x_2 + 2(1-\alpha_1)x_1 = 0 \\ \alpha_1 x_1 + (1-\alpha_1)(4x_2 - 2) = 0 \\ 0 \leq \alpha_1 \leq 1, \end{cases}$$

We obtain:

$$\begin{cases} x_1 = \dfrac{2\alpha_1(1-\alpha_1)}{-7\alpha_1^2 + 16\alpha_1 - 8}, \\ x_2 = \dfrac{4(1-\alpha_1)^2}{7\alpha_1^2 - 16\alpha_1 + 8} \\ 0 \leq \alpha_1 \leq 1, \\ \alpha_1 \neq \frac{8-2\sqrt{2}}{7} \end{cases} \tag{3}$$

In particular, for $\alpha_1 = \frac{8-2\sqrt{2}}{7}$ the previous system is impossible.

We obtain the set of points

$$\begin{cases} x_1 = \dfrac{2\alpha_1(1-\alpha_1)}{-7\alpha_1^2 + 16\alpha_1 - 8}, \\ x_2 = \dfrac{4(1-\alpha_1)^2}{7\alpha_1^2 - 16\alpha_1 + 8} \\ 0 \leq \alpha_1 < \frac{8-2\sqrt{2}}{7} \end{cases}$$

which are Pareto minima.

For $1 \geq \alpha_1 > \frac{8-2\sqrt{2}}{7}$ is not a minimum of (P_{α_1}) being the Hessian matrix Q indefinite as can also be checked by the previous Matlab program.

4) Consider the following matrix game:

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

(a) Find the dominated strategies 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.

SOLUTION

(a) Strategy 1 of Player 2 is dominated by Strategy 4, so that column 1 can be deleted. The reduced game is given by the matrix

$$C_{R1} = \begin{pmatrix} 4 & -1 & 5 & 2 \\ 1 & 3 & 3 & 5 \\ 3 & -2 & 3 & 1 \\ 1 & 5 & 2 & 3 \end{pmatrix}$$

Now, Strategy 1 of Player 1 is dominated by Strategy 3, so that row 1 can be deleted and the reduced matrix becomes:

$$C_{R2} = \begin{pmatrix} 1 & 3 & 3 & 5 \\ 3 & -2 & 3 & 1 \\ 1 & 5 & 2 & 3 \end{pmatrix}$$

(b) We observe that no minimum component on the columns of the reduced matrix is a maximum on the respective row, so that no pure strategies Nash equilibrium exists.

(c) The optimization problem associated with Player 1 is

$$\begin{cases} \min\ v \\ v \geq x_1 + 2x_2 + 2x_3 + x_4 \\ v \geq 4x_1 + x_2 + x_3 + x_4 \\ v \geq -x_1 + 3x_2 - 2x_3 + 5x_4 \\ v \geq 5x_1 + 3x_2 + 3x_3 + 2x_4 \\ v \geq 2x_1 + 5x_2 + x_3 + 3x_4 \\ x_1 + x_2 + x_3 + x_4 = 1 \\ x \geq 0 \end{cases} \tag{4}$$

The previous problem can be solved by Matlab.

Matlab solution

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

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 $(x, v) = (0, 0, \frac{3}{8}, \frac{5}{8}, 2.375)$. The optimal solution of the dual of (4) is given by $(y, w) = (0, 0, \frac{1}{8}, \frac{7}{8}, 0, 2.375)$. y can be found in the vector *lambda.ineqlin* given by the Matlab function linprog.

Therefore,

$$(x_1, x_2, x_3, x_4) = (0, 0, \frac{3}{8}, \frac{5}{8}), \quad (y_1, y_2, y_3, y_4, y_5) = (0, 0, \frac{1}{8}, \frac{7}{8}, 0),$$

is a mixed strategies Nash equilibrium.