

Πίνακας 9.1: Παράμετροι Αλγορίθμου

a	γ	η	ρ	ζ
0.4	0.1	1.5	0.5	5 %

Κεφάλαιο 9

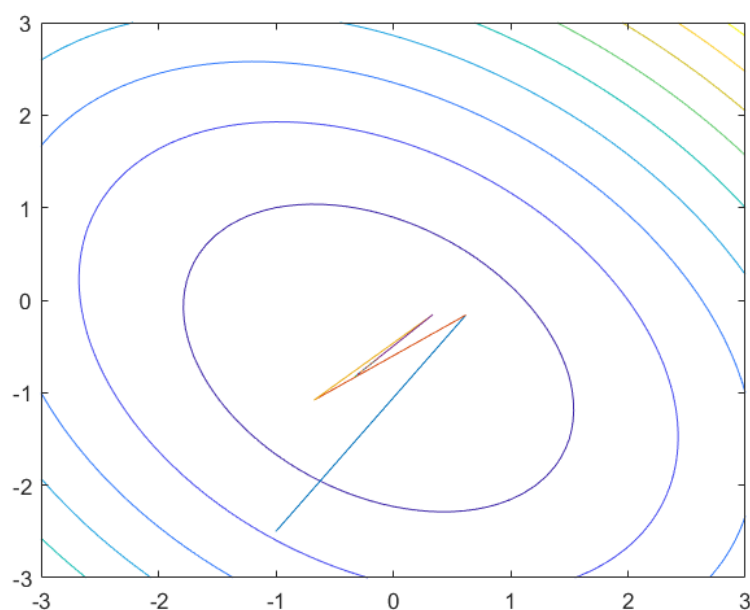
Πρόβλημα 9

Στο πρόβλημα αυτό υλοποιούνται 3 επαναλήψεις του variable learning rate αλγορίθμου, για την συνάρτηση 8.1 του Προβλήματος 8 .

Οι παράμετροι του αλγορίθμου που χρησιμοποιούνται αναγράφονται στον πίνακα 9.1, ενώ η αρχική εκτίμηση είναι :

$$x_0 = \begin{bmatrix} -1 \\ -2.5 \end{bmatrix} \quad (9.1)$$

$$\nabla F = Ax + D = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3x_1 + x_2 + 1 \\ x_1 + 3x_2 + 2 \end{bmatrix} \quad (9.2)$$



Εικόνα 9.1: Plot algorithm trajectory on contour plot of $F(x)$

Πίνακας 9.2: Δεδομένα εξόδου 1^{ης} επανάληψης

x_1	Διορθωμένο x_1	$F(x_1)$	Διορθωμένο $F(x_1)$	Διορθωμένο a	Διορθωμένο γ
0.62 -0.16	-1 -2.5	2.8158	9.3750	0.2	0

9.1 1^η Επανάληψη

Στο σημείο x_0 :

$$g_0 = \nabla F(x_0) = \begin{bmatrix} 3 \cdot -1 + -2.5 + 1 \\ -1 + 3 \cdot -2.5 + 2 \end{bmatrix} = \begin{bmatrix} -4.5 \\ -6.5 \end{bmatrix} \quad (9.3)$$

$$\Delta x_0 = \gamma \cdot \Delta x_1 - (1 - \gamma) \cdot a \cdot g_0 = 0.1 \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} - (1 - 0.1) \cdot 0.4 \cdot \begin{bmatrix} -4.5 \\ -6.5 \end{bmatrix} = \begin{bmatrix} -0.9 \cdot 0.4 \cdot -4.5 \\ -0.9 \cdot 0.4 \cdot -6.5 \end{bmatrix} = \begin{bmatrix} 1.62 \\ 2.34 \end{bmatrix} \quad (9.4)$$

$$x_{1new} = x_0 + \Delta x_0 = \begin{bmatrix} -1 \\ -2.5 \end{bmatrix} + \begin{bmatrix} 1.62 \\ 2.34 \end{bmatrix} = \begin{bmatrix} 0.62 \\ -0.16 \end{bmatrix} \quad (9.5)$$

Για να επαληθεύσουμε την εγκυρότητα του βήματος του αλγορίθμου, υπολογίζουμε την τιμή της συνάρτησης στο νέο αυτό σημείο x_1 . $F(x_1) = 2.8158$ Παρατηρούμε πως η τιμή της συνάρτησης είναι μεγαλύτερη της $F(x_0) = 9.3750$ κατά μεγαλύτερο ποσοστό της παραμέτρου μας $\zeta = 5\%$, για την ακρίβεια είναι τόσο % μεγαλύτερη. Συνεπώς, το δοκιμαστικό αυτό βήμα απορρίπτεται και το a μειώνεται, ενώ ο συντελεστής ορμής τίθεται ίσος με μηδέν. Τα δεδομένα εξόδου της 1^{ης} επανάληψης αναγράφονται αναλυτικά στον πίνακα 9.5, ενώ ο υπολογισμός αυτών γίνεται κάτωθεν του προαναφερθέντος πίνακα.

1.

$$x_1 = x_0 \quad (9.6)$$

2.

$$F(x_1) = F(x_0) = 2.8158 \quad (9.7)$$

3.

$$a = \rho \cdot a = 0.5 \cdot 0.4 = 0.2 \quad (9.8)$$

4.

$$\gamma = 0 \quad (9.9)$$

9.2 2^η Επανάληψη

Στο σημείο x_1 :

$$\Delta x_1 = -a \cdot g_1 = -0.2 \cdot -g_0 = -0.2 \cdot \begin{bmatrix} -4.5 \\ -6.5 \end{bmatrix} = \begin{bmatrix} 0.9 \\ 1.3 \end{bmatrix} \quad (9.10)$$

Πίνακας 9.3: Δεδομένα εξόδου 2^{ης} επανάληψης

x_2	Διορθωμένο x_2	$F(x_2)$	Διορθωμένο $F(x_2)$	Διορθωμένο a	Διορθωμένο γ
$\begin{bmatrix} -0.1 \\ -1.2 \end{bmatrix}$	$\begin{bmatrix} -1 \\ -2.5 \end{bmatrix}$	1.7950	9.3750	0.1	0

$$x_{2new} = x_1 + \Delta x_1 = \begin{bmatrix} -1 \\ -2.5 \end{bmatrix} + \begin{bmatrix} 0.9 \\ 1.3 \end{bmatrix} = \begin{bmatrix} -0.1 \\ -1.2 \end{bmatrix} \quad (9.11)$$

Για να επαληθεύσουμε την εγκυρότητα του βήματος του αλγορίθμου, υπολογίζουμε την τιμή της συνάρτησης στο νέο αυτό σημείο x_2 . $F(x_2) = 1.7950$ Παρατηρούμε πως η τιμή της συνάρτησης είναι μεγαλύτερη της $F(x_1) = 2.8158$ κατά μεγαλύτερο ποσοστό της παραμέτρου μας $\zeta = 5\%$, για την ακρίβεια είναι τόσο % μεγαλύτερη. Συνεπώς, και αυτό το δοκιμαστικό αυτό βήμα απορρίπτεται και το a μειώνεται, ενώ ο συντελεστής ορμής τίθεται ίσος με μηδέν. Τα δεδομένα εξόδου της 1^{ης} επανάληψης αναγράφονται αναλυτικά στον πίνακα 9.3, ενώ ο υπολογισμός αυτών γίνεται κάτωθεν του προαναφερθέντος πίνακα.

1.

$$x_2 = x_1 \quad (9.12)$$

2.

$$F(x_2) = F(x_1) = 9.3750 \quad (9.13)$$

3.

$$a = \rho \cdot a = 0.5 \cdot 0.2 = 0.1 \quad (9.14)$$

4.

$$\gamma = 0 \quad (9.15)$$

9.3 3^η Επανάληψη

Στο σημείο x_2 :

$$\Delta x_2 = -a \cdot g_2 = -0.1 \cdot -g_0 = -0.1 \cdot \begin{bmatrix} -4.5 \\ -6.5 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.65 \end{bmatrix} \quad (9.16)$$

$$x_{3new} = x_2 + \Delta x_2 = \begin{bmatrix} -1 \\ -2.5 \end{bmatrix} + \begin{bmatrix} 0.45 \\ 1.65 \end{bmatrix} = \begin{bmatrix} -0.55 \\ -0.85 \end{bmatrix} \quad (9.17)$$

Για να επαληθεύσουμε την εγκυρότητα του βήματος του αλγορίθμου, υπολογίζουμε την τιμή της συνάρτησης στο νέο αυτό σημείο x_2 . $F(x_2) = 1.7550$ Παρατηρούμε πως η τιμή της συνάρτησης **δεν** είναι μεγαλύτερη της $F(x_1) = 2.8158$ κατά μεγαλύτερο ποσοστό της παραμέτρου μας $\zeta = 5\%$, για την ακρίβεια είναι τόσο % μικρότερη. Συνεπώς, αυτό το δοκιμαστικό αυτό βήμα είναι **αποδεκτό** και το a αυξάνεται, ενώ ο συντελεστής ορμής τίθεται ίσος με την αρχική του τιμή. Τα δεδομένα εξόδου της 1^{ης} επανάληψης αναγράφονται

Πίνακας 9.4: Δεδομένα εξόδου 3^{ης} επανάληψης

x_3	$F(x_3)$	Διορθωμένο a	Διορθωμένο γ
$\begin{bmatrix} -0.55 \\ -0.85 \end{bmatrix}$	1.7550	0.15	0.1

Πίνακας 9.5: Δεδομένα εξόδου αλγορίθμου

Επανάληψη	x_k	Διορθωμένο x_k	$F(x_k)$	Διορθωμένο $F(x_1)$	Διορθωμένο a	Διορθωμένο γ
1 ^η	$\begin{bmatrix} 0.62 \\ -0.16 \end{bmatrix}$	$\begin{bmatrix} -1 \\ -2.5 \end{bmatrix}$	2.8158	9.3750	0.2	0
2 ^η	$\begin{bmatrix} -0.1 \\ -1.2 \end{bmatrix}$	$\begin{bmatrix} -1 \\ -2.5 \end{bmatrix}$	1.7950	9.3750	0.1	0
3 ^η	$\begin{bmatrix} -0.55 \\ -0.85 \end{bmatrix}$	$\begin{bmatrix} -0.55 \\ -0.85 \end{bmatrix}$	1.7550	1.7550	0.15	0.1

αναλυτικά στον πίνακα 9.4, ενώ ο υπολογισμός αυτών γίνεται άνωθεν του προαναφερθέντος πίνακα.

1.

$$x_3 = x_{3new} = \begin{bmatrix} -0.55 \\ -0.85 \end{bmatrix} \quad (9.18)$$

2.

$$F(x_3) = 1.7550 \quad (9.19)$$

3.

$$a = \eta \cdot a = 1.5 \cdot 0.1 = 0.15 \quad (9.20)$$

4.

$$\gamma = 0 \quad (9.21)$$

σεξτιονMatlab Code

```

1 x0 = [-1; -2.5];
2 A = [3 1; 1 3];
3 d = [1 2];
4
5 F_x0 = 0.5 * x0' * A * x0 + d * x0 + 2
6
7 x1t = [0.62; -0.16];
8
9 F_x1t = 0.5 * x1t' * A * x1t + d * x1t + 2
10
11 x2t = [-0.1; -1.2];
12
13 F_x2t = 0.5 * x2t' * A * x2t + d * x2t + 2
14
15 x3t = [-0.55; -0.85];
16
17 F_x3t = 0.5 * x3t' * A * x3t + d * x3t + 2

```

```

1 clear
2 [x1 x2] = meshgrid(-3:0.1:3);
3 z = 1.5*(x1.^2) + (x1.*x2) + 1.5*(x2.^2) + 1.*x1 + 2.*x2 + 2;
4 figure;
5 contour(x1, x2, z);
6 hold on;
7
8 alfa = 0.4;
9 gamma_0 = 0.1;
10 eta = 1.5;
11 rou = 0.5;
12 ksi = 0.05;
13
14 x0 = [-1 -2.5]';
15 A = [3 1;1 3];
16 d = [1 2]';
17 fold = (0.5)*x0'*A*x0 + d'*x0;
18 dx = zeros(2,1);
19 xold = x0;
20 g0 = A*xold + d;
21
22 z = [0 0];
23 w = [0 0];
24 while abs(g0) > 0.001
25     dx = gamma_0*dx - (1-gamma_0)*alfa*g0;
26     xt = xold + dx;
27     ft = (0.5)*xt'*A*xt + d'*xt;
28     df = (ft - fold)/fold;
29     if df > ksi
30         alfa = alfa * rou;
31         gamma = 0;
32     elseif df <= ksi & df > 0
33         z = [xold(1) xt(1)];
34         w = [xold(2) xt(2)];
35         plot(z,w);
36         xold = xt;
37         fold = ft;
38     else
39         z = [xold(1) xt(1)];
40         w = [xold(2) xt(2)];
41         plot(z,w);
42         xold = xt;
43         fold = ft;
44         alfa = alfa * eta;
45         gamma = 0.1;
46     end
47     g0 = A*xold + d;
48 end

```