23MAT204 – Mathematics for Intelligent Systems - 3

Practise Sheet-5

(Gradient Descent Method)

Gradient Descent Method to solve –

- **(i)** **Linear System Ax=b, where A^T=A**

  or the equivalent optimization problem

- **(ii)** **Minimize f(x), where f(x) is a quadratic optimization problem:**

$$Let \;\; f(x) = \frac{1}{2}x^T A x - b^T x + c, \;\; x \in R^n, A = A^T, c \in R$$

$Start$ with arbitrary x_0

$\nabla f(x) = Ax - b \Rightarrow \; Gradient \; at \; x = x_0$ is $Ax_0 - b$

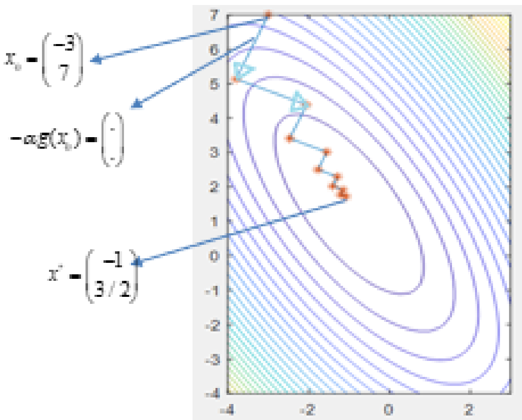
We denote this as $g(x_0)$

$Negative$ gradient is $r_0 = b - Ax_0$. r_0 is called residual

New updated $x \; is \;\; x_1 = x_0 - \alpha \, g(x_0)$

$What \; is$ a good α ?

Solution



$Let \;\; f(x) = \frac{1}{2}x^T A x - b^T x + c, \;\; x \in R^n, A = A^T, \;\; c \in R$

$Start$ with arbitrary x_0 (position vector)

$\nabla f(x) = Ax - b \Rightarrow \; Gradient \; at \; x = x_0$ is $g(x_0) = Ax_0 - b$

$Note$: Never forget that g(.) is a vector (displacement vector)

$Let \; x_1 = x_0 - \alpha \, g(x_0). \;\; We \; donot \; know \; \alpha \; to \; compute \; x_1$

Let us find it through optimization.

$$f(x_0 - \alpha\, g(x_0)) = \frac{1}{2}(x_0 - \alpha\, g(x_0))^T A (x_0 - \alpha\, g(x_0)) - b^T (x_0 - \alpha\, g(x_0)) + c$$

$$= -\alpha x_0^T A g(x_0) + \frac{\alpha^2}{2}(g(x_0))^T A g(x_0) + \alpha b^T \, g(x_0) + K, \quad \text{where K is constant}$$

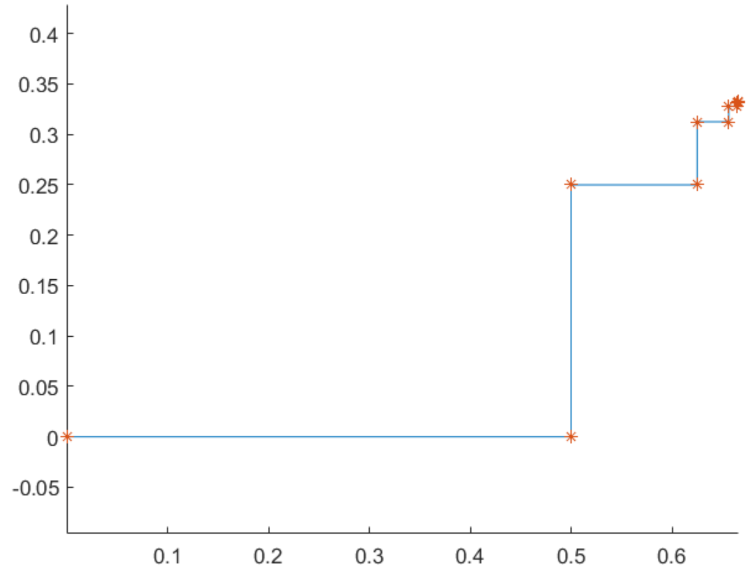
$$\frac{df}{d\alpha} = -x_0^T A g(x_0) + b^T \, g(x_0) + \alpha \, (g(x_0))^T A g(x_0)$$

$$\frac{df}{d\alpha} \Rightarrow 0 \Rightarrow \alpha = \frac{(g(x_0))^T (Ax_0) - (g(x_0))^T b}{(g(x_0))^T A g(x_0)} = \frac{(g(x_0))^T (Ax_0 - b)}{(g(x_0))^T A g(x_0)} = \frac{(g(x_0))^T g(x_0)}{(g(x_0))^T A g(x_0)}$$

Example 1:

Solve the system: $2x_1-x_2=1$; $-x_1+2x_2=0$; using gradient descent algorithm with initial vector as $(0,0)^T$

```
A=[2,-1;-1,2]; b=[1;0];
xo=[0;0]; % starting point, column vector
xa=xo; % store in array
for i=1:10
g=A*xo-b;
alpha=g'*g/(g'*A*g);
xn=xo-alpha*g;
xa=[xa xn];
xo=xn;
errnorm=norm(A*xn-b,2);
if errnorm <0.001
break;
end
end
hold on
plot(xa(1,:),xa(2,:))
hold on
plot(xa(1,:),xa(2,:),'*')
axis equal
xa
```

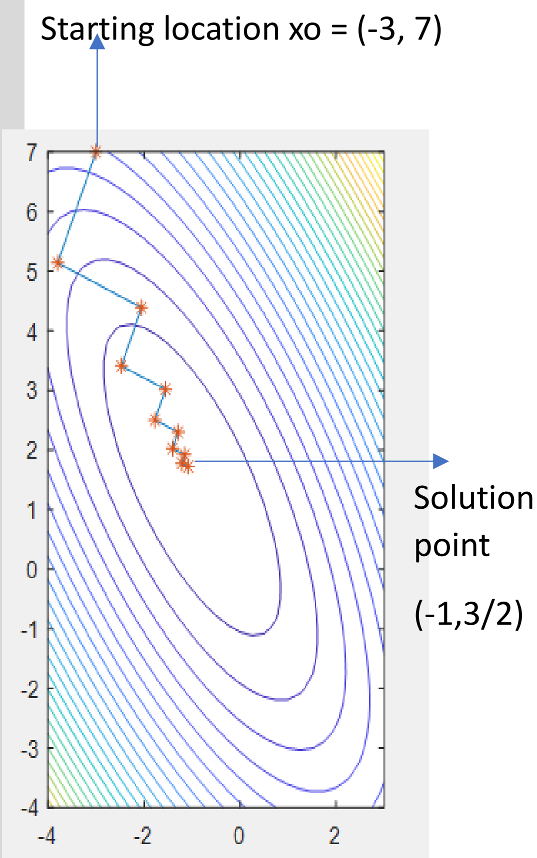



```
xa = 2×11
         0    0.5000    0.5000    0.6250    0.6250    0.6563    0.6563    0.6641    0.6641    0.6660    0.6660
         0         0    0.2500    0.2500    0.3125    0.3125    0.3281    0.3281    0.3320    0.3320    0.3330
```

Example 2:

Find the minimum of $f(x)=2x_1^2+ x_1^2+2\, x_1\, x_2+ x_1- x_2$ using gradient descent algorithm

```
x1=-4:.25:3;
x2=-4:.25:7;
[X1,X2] = meshgrid(x1,x2);
Z = 2*X1.^2+X2.^2+2*X1.*X2+X1-X2;
contour(X1,X2,Z,30);
xo=[-3;7]; % starting point, column vector
A=[4 2; 2 2]; b=[-1;1];
xa=xo; % store in array
for i=1:10
     g=A*xo-b;
     alpha=g'*g/(g'*A*g);
     xn=xo-alpha*g;
     xa=[xa xn];
     xo=xn;
     errnorm=norm(A*xn-b,2);
     if errnorm <0.001
         break;
     end
end
hold on
 plot(xa(1,:),xa(2,:))
 hold on
 plot(xa(1,:),xa(2,:),'*')
 axis equal
```



Practice questions:

1. Solve the given linear systems using gradient descent method by taking different starting points.
   (a) 2x+y-z=1, x+2y-z = 2, -x-y+4z=9
   (b) Ax=b, where $A = \begin{bmatrix} 11 & 2 & 1 \\ 2 & 4 & -1 \\ 1 & -1 & 6 \end{bmatrix}$ and $b = \begin{bmatrix} 12 \\ 1 \\ 7 \end{bmatrix}$
   (c) Ax=b, where $A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{pmatrix}, b = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$

2. Consider the optimization problem:
   Minimize $f(x,y) = 4x^2 + 3y^2 - 16x - 36y + 25$
   (a) Solve the problem analytically and obtain the solution.
   (b) Starting from (0,0), perform gradient descent method and obtain the solution. Also plot the path taken to reach the final solution.
   (c) Starting from (50,0), perform gradient descent method and obtain the solution. Also plot the path taken to reach the final solution.
   (d) Starting from (0,-25), perform gradient descent method and obtain the solution. Also plot the path taken to reach the final solution.
   (e) Starting from (-21,35), perform gradient descent method and obtain the solution. Also plot the path taken to reach the final solution.

3. Solve the given quadratic optimization problems using gradient descent method by taking different starting points.
   (a) Minimize $f(x,y,z) = 6x^2 + 8y^2 + z^2 + 2xz + 4yz - 3x - 3z$
   (b) Minimize $f(x,y,z) = 9x^2 + 5y^2 + 3z^2 - 36x + 30y - 8z$

4. Solve the system AX=B, where $A = \begin{bmatrix} 9 & 1 & 1 & 0 & 0 \\ 1 & 5 & 0 & 0 & 2 \\ 1 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 2 \\ 0 & 2 & 0 & 2 & 6 \end{bmatrix}$ and $\begin{bmatrix} 10 \\ 3 \\ 4 \\ 2 \\ 6 \end{bmatrix}$ using gradient descent method with initial vector as x=[a,b,c,b,a]T, where: a is the last two digits of your registration number, b is your date of birth, c is your month of birth.

5. Generate a random integer squaresymmetric matrix A of order 5. Obtain the vector b, such that Ax=b, with x=$[1,2,3,4,5,6,7,8,9]^T$. Solve the system AX=b using gradient descent method and verify the solution X=x. In how many iterations you could get the solution?