1 Order and Rate of Convergence

1.1 $x^{(k)} = 2 + 0.2^k$

Clearly, the algorithm converges to 2.

$$\lim_{k\to\infty} \frac{||x^{k+1} - x^*||}{||x^{(k)} - x^*||} = \lim_{k\to\infty} \frac{0.2^{k+1}}{0.2^k} = 0.2$$

The algorithm converge at least linearly.

$$\lim_{k\to\infty} \frac{||x^{(k+1)} - x^*||}{||x^{(k)} - x^*||^p} = \frac{0.2^{k+1}}{0.2^{kp}} = \frac{0.2}{0.2^{k(p-1)}} = \infty$$

Linear convergence.

1.2 $x^{(k)} = 1 + (\frac{4}{k})^k$

$$\lim_{k\to\infty} x^{(k)} = 1 + \lim_{k\to\infty} (\frac{4}{k})^k = 1$$

$$\lim_{x\to\infty} \frac{||x^{(k+1)} - x^*||}{||x^{(k)} - x^*||} = \frac{(4/(k+1))^{(k+1)}}{(4/k)^k} = 0$$

The algorithm converge at least linearly.

$$\lim_{k\to\infty} \frac{||x^{(k+1)} - x^*||}{||x^{(k)} - x^*||^p} = \frac{(4/(k+1))^{(k+1)}}{(4/k)^{pk}} = \infty$$

Linear convergence.

1.3 $x^{(k)} = 2^{-2^k}$

$$\lim_{k\to\infty} x^{(k)} = \lim_{k\to\infty} 2^{-2^k} = 0$$

$$\lim_{k\to\infty} \frac{||x^{(k+1)} - x^*||}{||x^{(k)} - x^*||} = \frac{2^{-2^{(k+1)}}}{2^{-2^k}} = 0$$

The algorithm converge at least linearly.

$$\lim_{k\to\infty} \frac{||x^{(k+1)} - x^*||}{||x^{(k)} - x^*||^p} = \frac{2^{-2^{(k+1)}}}{2^{-2^{(pk)}}} = \infty$$

Linear convergence.

1.4 $x^{(k+1)} = \frac{1}{2}x^{(k)} + \frac{8}{x^{(k)}}$

Suppose the algorithm converges, we have

$$x^* = \frac{1}{2}x^* + \frac{8}{x^*}$$

$$x^* = \pm 4$$

The problem gives $x^{(0)} \leq 0$, we have $x^* = -4$.

$$\lim_{k\to\infty} \frac{||x^{(k+1)} - x^*||}{||x^{(k)} - x^*||^2} = \lim_{k\to\infty} \frac{||\frac{1}{2}x^{(k)} + \frac{8}{x^{(k)}} + 4||}{||x^{(k)} + 4||^2} = \lim_{k\to\infty} \frac{1}{||2x^{(k)}||} = \frac{1}{8}$$

The algorithm converges at order 2.

2 Convergence of SD

$$f(\boldsymbol{x}) = x_1^2 + 2x_2^2$$

Starting at $\boldsymbol{x}^{(0)} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

2.1 Recurrence Relation

We verify that using numeric example for the first 5 iterations.

$$x^{(1-5)} = \begin{pmatrix} 2 & 0.6667 & 0.2222 & 0.0741 & 0.0247 \\ 1 & -0.3333 & 0.1111 & -0.0370 & 0.0123 \end{pmatrix}$$

$$x_k^{(1-5)} = \begin{pmatrix} 2 & 0.6667 & 0.2222 & 0.0741 & 0.0247 \\ 1 & -0.3333 & 0.1111 & -0.0370 & 0.0123 \end{pmatrix}$$

which are equivalent. Similarly,

$$f^{(1-4)}/9 = \begin{pmatrix} 0.6667 & 0.0741 & 0.0082 & 9.1111\text{e-}04 \end{pmatrix}$$

$$f^{(2-5)} = \begin{pmatrix} 0.6667 & 0.0741 & 0.0082 & 9.1449\text{e-}04 \end{pmatrix}$$

They are roughly equal. The error probably coming from round-off for division during the numeric computation.

See attached code for the validation of first 100 iteration.

2.2 $\boldsymbol{x}^*$ and Convergence Ratio

$$\lim_{k\to\infty} x^{(k)} = \lim_{k\to\infty} \frac{1}{3^k} \begin{pmatrix} 2 \\ (-1)^k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\lim_{x\to\infty} \frac{||x^{(k+1)} - x^*||}{||x^{(k)} - x^*||} = \lim_{x\to\infty} \frac{\left|\left| \begin{pmatrix} \frac{2}{3^{k+1}} \\ \frac{-1^{k+1}}{3^{k+1}} \end{pmatrix} \right|\right|}{\left|\left| \begin{pmatrix} \frac{2}{3^k} \\ \frac{-1^k}{3^k} \end{pmatrix} \right|\right|} = \frac{1}{3}$$

3 SD / NR

$$f(\boldsymbol{x}) = x_2(x_1 - 2)^2 + 10(x_2 - 1)^2$$

with $x^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

3.1 SD

- $k = 1$

$$d = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

$$x^{(1)} = \begin{pmatrix} 2\,\alpha + 1 \\ 1 - \alpha \end{pmatrix}$$

$$\phi = 10\,\alpha^2 - (2\,\alpha - 1)^2\,(\alpha - 1)$$

Solve for optimal α

$$\alpha = 0.1460$$

$$x^{(1)} = \begin{pmatrix} 1.2920 \\ 0.8540 \end{pmatrix}$$

- $k = 2$

$$d = \begin{pmatrix} 1.2093 \\ 2.4186 \end{pmatrix}$$

$$x^{(2)} = \begin{pmatrix} 1.2093\,\alpha + 1.2920 \\ 2.4186\,\alpha + 0.8540 \end{pmatrix}$$

$$\phi = 10\,(2.4186\,\alpha - 0.1460)^2 + (1.2093\,\alpha - 0.7080)^2\,(2.4186\,\alpha + 0.8540)$$

Solve for optimal α

$$\alpha = 0.0653$$

$$x^{(2)} = \begin{pmatrix} 1.3710 \\ 1.0120 \end{pmatrix}$$

3.2 NR

1. $k = 1$

$$H_f = \begin{pmatrix} 2 & -2 \\ -2 & 20 \end{pmatrix}$$

$$g = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$$d = \begin{pmatrix} 1.0556 \\ 0.0556 \end{pmatrix} ca$$

$$x^{(1)} = \begin{pmatrix} 1.0556\,\alpha + 1 \\ 0.0556\,\alpha + 1 \end{pmatrix}$$

$$\phi = (1.0556\,\alpha - 1)^2\,(0.0556\,\alpha + 1) + 0.0309\,\alpha^2$$

$$\alpha = 0.9230$$

$$x^{(1)} = \begin{pmatrix} 1.9743 \\ 1.0513 \end{pmatrix}$$

2. $k = 2$

$$H_f = \begin{pmatrix} 2.1026 & -0.0514 \\ -0.0514 & 20 \end{pmatrix}$$

$$g = \begin{pmatrix} -0.0540 \\ 1.0263 \end{pmatrix}$$

$$d = \begin{pmatrix} 0.0244 \\ -0.0512 \end{pmatrix} x^{(2)} = \begin{pmatrix} 0.0244\,\alpha + 1.9743 \\ 1.0513 - 0.0512\,\alpha \end{pmatrix}$$

$$\phi = 10\,(0.0512\,\alpha - 0.0513)^2 - (0.0512\,\alpha - 1.0513)\,(0.0244\,\alpha - 0.0257)^2$$

$$\alpha = 1.0017$$

$$x^{(2)} = \begin{pmatrix} 1.9988 \\ 0.9999 \end{pmatrix}$$

Prob 2 Code

```
syms x1 x2 alpha k
```

```
f = x1^2 + 2*x2^2
```

f = $x_1^2 + 2\,x_2^2$

```
g = gradient(f)
```

g =

$$\begin{pmatrix} 2\,x_1 \\ 4\,x_2 \end{pmatrix}$$

```
x = [x1;x2]
```

x =

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

```
x_val = [2;1]
```

```
x_val = 2×1
     2
     1
```

```
size = 100
```

```
size = 100
```

```
g_val = subs(g, x, x_val)
```

g_val =

$$\begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

```
all_x = zeros(2, size + 1)
```

```
all_x = 2×101
     0     0     0     0     0     0     0     0     0     0     0     0     0 ···
     0     0     0     0     0     0     0     0     0     0     0     0     0
```

```
all_f = zeros(1, size + 1)
```

```
all_f = 1×101
     0     0     0     0     0     0     0     0     0     0     0     0     0 ···
```

```
all_x(:, 1) = x_val
```

```
all_x = 2×101
     2     0     0     0     0     0     0     0     0     0     0     0     0 ···
```

```
     1     0     0     0     0     0     0     0     0     0     0     0     0
```

```
all_f(1) = subs(f, x, x_val)
```

```
all_f = 1×101
     6     0     0     0     0     0     0     0     0     0     0     0     0 ...
```

```
threshhold = 0.001
```

```
threshhold = 1.0000e-03
```

```
echo off;
for i = 2:size+1
    d = - g_val;
    x_new = x_val + alpha * d;

    phi = subs(f, x, x_new);

    alpha_vals = solve([gradient(phi) == 0, alpha >= 0], alpha);
    phi_vals = double(subs(phi, alpha, alpha_vals));
    [phi_vals, alpha_idx] = sort(phi_vals);

    alpha_val = alpha_vals(alpha_idx(1));

    x_val = subs(x_new, alpha, alpha_val);
    g_val = subs(g, x, x_val);

    all_x(:, i) = x_val;
    all_f(i) = subs(f, x, x_val);
%     norm(g_val), threshhold *
end
echo on;
```

```
syms k

xk = 1/3^k * [2; (-1)^k]
```

xk =

$$\begin{pmatrix} \frac{2}{3^k} \\ \frac{(-1)^k}{3^k} \end{pmatrix}$$

```
xks = double(subs(xk, 0:size))
```

```
xks = 2×101
    2.0000    0.6667    0.2222    0.0741    0.0247    0.0082    0.0027    0.0009 ...
    1.0000   -0.3333    0.1111   -0.0370    0.0123   -0.0041    0.0014   -0.0005
```

```
sym(xks(:, 1:5))
```

```
ans =
```

$$\begin{pmatrix} 2 & 0.6667 & 0.2222 & 0.0741 & 0.0247 \\ 1 & -0.3333 & 0.1111 & -0.0370 & 0.0123 \end{pmatrix}$$

```
sym(all_x(:, 1:5))
```

ans =

$$\begin{pmatrix} 2 & 0.6667 & 0.2222 & 0.0741 & 0.0247 \\ 1 & -0.3333 & 0.1111 & -0.0370 & 0.0123 \end{pmatrix}$$

```
% compare the computed x's from sd and the equation
all(all(all_x - xks < 0.0001))
```

ans = *logical*
   1

```
sym(all_f(1:4) / 9)
```

ans = $\begin{pmatrix} 0.6667 & 0.0741 & 0.0082 & 9.1111\text{e-}04 \end{pmatrix}$

```
sym(all_f(2:5))
```

ans = $\begin{pmatrix} 0.6667 & 0.0741 & 0.0082 & 9.1449\text{e-}04 \end{pmatrix}$

```
% compare the recurrence relation of f
all(all_f(1:size-1)/9 - all_f(2:size) < 0.0001)
```

ans = *logical*
   1

Prob 3 - SD Code

```
syms x [1 2]
syms alpha
```

```
f = x2 * (x1 - 2)^2 + 10 * (x2 - 1)^2
```

f = $x_2\ (x_1 - 2)^2 + 10\ (x_2 - 1)^2$

```
g = gradient(f, x)
```

g =

$$\begin{pmatrix} x_2\ (2\,x_1 - 4) \\ 20\,x_2 + (x_1 - 2)^2 - 20 \end{pmatrix}$$

```
x = transpose(x)
```

x =

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

```
x_val = [1;1]
```

```
x_val = 2×1
     1
     1
```

```
g_val = subs(g, x, x_val)
```

g_val =

$$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

```
threshhold = 0.001
```

threshhold = 1.0000e-03

```
for i = 1:2
    d = - g_val
    x_new = x_val + alpha * d

    phi = subs(f, x, x_new)

    alpha_vals = solve([gradient(phi) == 0, alpha >= 0], alpha)
    phi_vals = double(subs(phi, alpha, alpha_vals))
    [phi_vals, alpha_idx] = sort(phi_vals)

    alpha_val = alpha_vals(alpha_idx(1))

    x_val = subs(x_new, alpha, alpha_val)
```

```
    g_val = subs(g, x, x_val)

end
```

d =

$$\begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

x_new =

$$\begin{pmatrix} 2\,\alpha + 1 \\ 1 - \alpha \end{pmatrix}$$

phi = $10\,\alpha^2 - (2\,\alpha - 1)^2\;(\alpha - 1)$

alpha_vals =

$$\begin{pmatrix} 0.1460 \\ 2.8540 \end{pmatrix}$$

```
phi_vals = 2×1
    0.6412
   40.3588
phi_vals = 2×1
    0.6412
   40.3588
alpha_idx = 2×1
     1
     2
```

alpha_val = 0.1460

x_val =

$$\begin{pmatrix} 1.2920 \\ 0.8540 \end{pmatrix}$$

g_val =

$$\begin{pmatrix} -1.2093 \\ -2.4186 \end{pmatrix}$$

d =

$$\begin{pmatrix} 1.2093 \\ 2.4186 \end{pmatrix}$$

x_new =

$$\begin{pmatrix} 1.2093\,\alpha + 1.2920 \\ 2.4186\,\alpha + 0.8540 \end{pmatrix}$$

phi = $10\;(2.4186\,\alpha - 0.1460)^2 + (1.2093\,\alpha - 0.7080)^2\;(2.4186\,\alpha + 0.8540)$

alpha_vals = 0.0653

```
phi_vals = 0.4018
phi_vals = 0.4018
alpha_idx = 1
```

alpha_val = 0.0653

x_val =

$$\begin{pmatrix} 1.3710 \\ 1.0120 \end{pmatrix}$$

g_val =

$$\begin{pmatrix} -1.2731 \\ 0.6366 \end{pmatrix}$$

Prob 3 - NR Code

```
syms x [1 2]
syms alpha
```

```
f = x2 * (x1 - 2)^2 + 10 * (x2 - 1)^2
```

f = $x_2\,(x_1-2)^2+10\,(x_2-1)^2$

```
g = gradient(f, x)
```

g =

$$\begin{pmatrix} x_2\,(2\,x_1-4) \\ 20\,x_2+(x_1-2)^2-20 \end{pmatrix}$$

```
h = hessian(f)
```

h =

$$\begin{pmatrix} 2\,x_2 & 2\,x_1-4 \\ 2\,x_1-4 & 20 \end{pmatrix}$$

```
x = transpose(x)
```

x =

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

```
x_val = [1;1]
```

```
x_val = 2×1
     1
     1
```

```
for i = 1:2
    h_val = subs(h, x, x_val)
    g_val = subs(g, x, x_val)

    d_val = - inv(h_val) * g_val

    x_new = x_val + alpha * d_val

    phi = subs(f, x, x_new)

    alpha_vals = solve([gradient(phi) == 0, alpha >= 0], alpha)
    phi_vals = double(subs(phi, alpha, alpha_vals))
    [phi_vals, alpha_idx] = sort(phi_vals)

    alpha_val = alpha_vals(alpha_idx(1))
```

```
    x_val = subs(x_new, alpha, alpha_val)
    g_val = subs(g, x, x_val)
end
```

h_val =

$$\begin{pmatrix} 2 & -2 \\ -2 & 20 \end{pmatrix}$$

g_val =

$$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

d_val =

$$\begin{pmatrix} 1.0556 \\ 0.0556 \end{pmatrix}$$

x_new =

$$\begin{pmatrix} 1.0556\,\alpha + 1 \\ 0.0556\,\alpha + 1 \end{pmatrix}$$

phi = $(1.0556\,\alpha - 1)^2\ (0.0556\,\alpha + 1) + 0.0309\,\alpha^2$

alpha_vals = 0.9230

phi_vals = 0.0270
phi_vals = 0.0270
alpha_idx = 1

alpha_val = 0.9230

x_val =

$$\begin{pmatrix} 1.9743 \\ 1.0513 \end{pmatrix}$$

g_val =

$$\begin{pmatrix} -0.0540 \\ 1.0263 \end{pmatrix}$$

h_val =

$$\begin{pmatrix} 2.1026 & -0.0514 \\ -0.0514 & 20 \end{pmatrix}$$

g_val =

$$\begin{pmatrix} -0.0540 \\ 1.0263 \end{pmatrix}$$

d_val =

$$\begin{pmatrix} 0.0244 \\ -0.0512 \end{pmatrix}$$

x_new =

$$\begin{pmatrix} 0.0244\,\alpha + 1.9743 \\ 1.0513 - 0.0512\,\alpha \end{pmatrix}$$

phi = $10\ (0.0512\,\alpha - 0.0513)^2 - (0.0512\,\alpha - 1.0513)\ (0.0244\,\alpha - 0.0257)^2$

alpha_vals =

$$\begin{pmatrix} 586.2227 \\ 1.0017 \end{pmatrix}$$

phi_vals = 2×1

```
10^3 ×
    3.0670
    0.0000
phi_vals = 2×1
10^3 ×
    0.0000
    3.0670
alpha_idx = 2×1
     2
     1
```

alpha_val = 1.0017

x_val =

$$\begin{pmatrix} 1.9988 \\ 0.9999 \end{pmatrix}$$

g_val =

$$\begin{pmatrix} -0.0024 \\ -0.0012 \end{pmatrix}$$