ELEC 405 - Homework 1

Gul Sena Altintas - 64284

November 6, 2020

Contents

1. Sampling and Matrix Notation

Let $x \in \mathcal{R}^n$, ie. $x = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}$

1. Let $\mathbf{y} \in \mathcal{R}^{2n-1}$ and

$$y_k = \begin{cases} x_{\frac{k+1}{2}} & \text{if k is odd} \\ 0 & otherwise \end{cases} \tag{1}$$

for $k = 1, \dots, 2n-1$ and $y = \begin{bmatrix} y_1 & y_2 & \dots & y_{2n-1} \end{bmatrix}^T$.

Answer: y^T is of the form $\begin{bmatrix} x_1 & 0 & x_2 & 0 & \dots & x_n \end{bmatrix}$. So, $A \in \mathcal{R}^{2n-1 \times n}$ and it should be of the form:

$$A_{ij} = \begin{cases} 1 & \text{if i is odd and } j = \frac{i+1}{2} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

2. Assume n is even and $y \in \mathcal{R}^{\frac{n}{2}}$. Define $y_k = x_{2k}$ for $k = 1, \dots, \frac{n}{2}$.
Answer: $A \in \mathcal{R}^{\frac{n}{2} \times n}$ and

$$A_{ij} = \begin{cases} 1 & \text{if } j = 2 \times i \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

3. Assume n is even and $y \in \mathcal{R}^{\frac{n}{2}}$. Define $y_k = \frac{x_{2k} + x_{2k-1}}{2}$ for $k = 1, \dots, \frac{n}{2}$.
Answer: $A \in \mathcal{R}^{\frac{n}{2} \times n}$ and

$$A_{ij} = \begin{cases} \frac{1}{2} & \text{if } j = 2 \times i \text{ or } j = 2 \times i - 1 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

2. Vector Space of Polynomials

Let $\mathcal{P}_{n-1}$ represent the vectors space of polynomials with degree less than or equal to $n-1$. Therefore, each element in this vector space can be written as

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1}x^{n-1} \tag{5}$$

Answer: $b = \frac{dp(x)}{dx}, b \in \mathcal{R}^{n-1}$ and then $b = \begin{bmatrix} 1\times c_1 & 2\times c_2 & \dots & (n-1)\times c_{n-1} \end{bmatrix}^T$. Then

$$b = Ha$$

$H \in \mathcal{R}^{n-1\times n}$ and

$$H_{ij} = \begin{cases} i & \text{if } i = j+1 \\ 0 & \text{otherwise} \end{cases}$$
$$H = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & & \dots & & \\ & & & & n-1 \end{bmatrix}$$

3. Affine functions

1. Let $x, y \in \mathcal{R}^n$ and $\alpha + \beta = 1$.

$$\begin{aligned} f(\alpha x + \beta y) &= A(\alpha x + \beta y) + b \\ &= \alpha Ax + \beta Ay + \alpha b + (1-\alpha)b \\ &= \alpha Ax + \beta Ay + \alpha b + \beta b \\ &= \alpha(Ax + b) + \beta(Ax + b) \\ &= \alpha f(x) + \beta f(y) \end{aligned}$$

2. Let $f : \mathcal{R}^n \to \mathcal{R}^m$ and be an affine function and $g : \mathcal{R}^n \to \mathcal{R}^m$ such that $g(x) = f(x) - f(0)$. We will show that $g(x)$ is linear. For a function to be linear, it must satisfy two properties: additivity and homogeneity.

   - **Homogeneity:** $g(\alpha x) = \alpha g(x)$

$$\begin{aligned} g(\alpha x) &= f(\alpha x) - f(0) \\ &= f(\alpha x + (1-\alpha)0) - f(0) \\ &= \alpha f(x) + (1-\alpha)f(0) - f(0) \\ &= \alpha(f(x) - f(0)) \\ &= \alpha g(x) \end{aligned} \tag{6}$$

   Eq (6) follows from the fact that f is affine. Therefore, g satisfies the homogeneity property.

   - **Addivity:** $g(x + y) = g(x) + g(y)$

$$\begin{aligned} g(x+y) &= f(x+y) - f(0) \\ &= f(\frac{1}{2}(2x) + \frac{1}{2}(2y)) - f(0) \\ &= \frac{1}{2}(f(2x) - f(0)) + \frac{1}{2}(f(2y) - f(0)) \end{aligned} \tag{7}$$

$$= \frac{1}{2}g(2x) + \frac{1}{2}g(2y) \tag{8}$$
$$= g(x) + g(y)$$

Eq (7) follows from the fact that f is affine. Eq(8) uses the homogeneity property of g proved earlier.

Therefore, g is linear, in other words $\exists A \in \mathcal{R}^{m\times n}$ such that $g(x) = Ax$ for any $x \in \mathcal{R}^n$. Then $f(x)$ can be rewritten as

$$f(x) = g(x) + f(0) = Ax + b$$

where $b = f(0)$ and $b \in \mathcal{R}$.

4. Yet Another Proof of Cauchy-Schwarz...

1. Let $a \geq 0, c \geq 0$ and $\forall\lambda \in \mathcal{R}, a + 2b\lambda + c\lambda^2 \geq 0$.
   Let $f : \mathcal{R} \to \mathcal{R}$ and $f(\lambda) = a + 2b\lambda + c\lambda^2$. For f to be ≥ 0 for all $\lambda \in \mathcal{R}$, f should have one or zero real roots otherwise we can find $\lambda' \in \mathcal{R}$ such that $f(\lambda') < 0$ (Easy to visualize geometrically). So, the determinant Δ should be ≤ 0.

$$\Delta = (2b)^2 - 4ac \leq 0$$
$$b^2 \leq ac$$
$$|b| \leq \sqrt{ac} \tag{9}$$

   (9) holds because $a, c \geq 0$.

2. Given $u, w \in \mathcal{R}^n$, explain why $(u + \lambda w)^T(u + \lambda w) \geq 0 \ \forall\lambda \in \mathcal{R}$.

$$\begin{aligned}(u + \lambda w)^T(u + \lambda w) &= < u + \lambda w, u + \lambda w > \\ &= < u, u + \lambda w > + < \lambda w, u + \lambda w > \\ &= < u, u > + < u, \lambda w > + < \lambda w, u > + < \lambda w, \lambda w > \\ &= \|u\|^2 + 2\lambda u^T w + \lambda^2\|w\|^2 \qquad (10) \\ &= (u + \lambda w)^2 \\ &\geq 0\end{aligned}$$

   (10) and previous ones follow from the properties of inner products.

3. Eq (10) is of the form of quadratic equation in the first part. Namely, here $a = \|u\|^2, c = \|w\|^2$ and $b =< u, w >$ $a, c \geq 0$ for any $\lambda \in \mathcal{R}$. Then by the first part, we know that $|b| \leq \sqrt{ac}$ so

$$\begin{aligned}|u^T w| &\leq \sqrt{\|u\|^2\|w\|^2} \\ &\leq \sqrt{u^T u} \times \sqrt{w^T w} \qquad (11)\end{aligned}$$

   (11) follows from the fact that $\|u\|^2 =< u, u >= u^T u$.

5. Inequality Proof

Let $a, b \in \mathcal{R}^n$, and $x = a + b, y = -b$ by triangle inequality

$$\|x + y\| \leq \|x\| + \|y\|$$
$$\|a\| \leq \|a + b\| + \| - b\|$$
$$\|a\| - \|b\| \leq \|a + b\|$$

6. First steps in multivariable quadratic functions

$$f(x_1, x_2, x_3) = 2x_1^2 + 3x_1x_3 - 2x_3x_2 - x_2^2 + 4x_3^2$$

can be written of the form x^TAx and A is not unique because of the non-zero x_{ij} terms. As long as $A_{ij} + A_{ji}$ equals to the coefficient of $x_i \times x_j$. Two possible choices of A is $\begin{bmatrix} 2 & 0 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & 4 \end{bmatrix}$ and its transpose $\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 3 & -2 & 4 \end{bmatrix}$.

7. On Gradients....

Let $f : \mathcal{R}^n \to \mathcal{R}$ be a differentiable multivariate function.

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \tag{12}$$

Find the explicit gradients of the following fuunctions

1. $f(x) = a^Tx + b$ for some $a \in \mathcal{R}^n, b \in \mathcal{R}$.
   $\nabla f(x) = a$ because $f(x) = a_1x_1 + x_2x_2 + \dots a_nx_n + b$ so $\frac{\partial f}{\partial x_i} = a_i$.

2. $f(x) = x^TAx$ for some $A \in \mathcal{R}^{n\times n}$.
   $\nabla f(x) = (A + A^T)x$ because $f(x) = x^TAx$ equivalently

$$\begin{aligned} f(x) &= x^TAx \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} x_iA_{ij}x_j \\ &= \sum_{i=1}^{n} x_iA_{i1}x_1 + \sum_{j=1}^{n} x_1A_{1j}x_j + \sum_{i=2}^{n}\sum_{j=2}^{n} x_iA_{ij}x_j & (13) \\ \frac{\partial f}{\partial x_1} &= \sum_{i=1}^{n} x_iA_{i1} + \sum_{j=1}^{n} A_{1j}x_j & (14) \end{aligned}$$

Let $A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}$ where $a_i, b_i \in \mathcal{R}^n$. Then (14) can be rewritten of the form $x^Tb_1 + x^Ta_1$. Note that x^Tb_1 corresponds to the first row of A^Tx so $\frac{\partial f}{\partial x_i} = (A_i + A_i^T)x_i^T$

8. What is a homework without a computational part?....

$f(x) = x^TAx \Rightarrow \nabla f(x) = 2Ax$, because A is symmetric.

Elec 405 - Homework 1

In [5]:
```
using LinearAlgebra
using Plots
```

8. Affine Approximation

$$a(\mathbf{x}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) \tag{7}$$

is an affine approximation of $f(\mathbf{x})$ around $\mathbf{x}_0$.
Note that this affine function defines an hyperplane in $\Re^{n+1}$ in the standard form

$$\begin{bmatrix} \nabla f(\mathbf{x}_0) \\ -1 \end{bmatrix}^T \left(\begin{bmatrix} \mathbf{x} \\ a(\mathbf{x}) \end{bmatrix} - \begin{bmatrix} \mathbf{x}_0 \\ f(\mathbf{x}_0) \end{bmatrix} \right) = 0. \tag{8}$$

Now consider the function

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} \tag{9}$$

defined over $\Re^2$, where

$$\mathbf{A} = \begin{bmatrix} 0.01 & 0.001 \\ 0.001 & 0.01 \end{bmatrix}. \tag{10}$$

In [6]:
```
A = [0.01 0.001; 0.001 0.01]
A
```

Out[6]:
```
2×2 Array{Float64,2}:
 0.01   0.001
 0.001  0.01
```

In [7]:
```
function draw_hyperplane_3D_modified(a, b, camera=(60, 45))

#     pyplot()
    x=range(-50, stop=50, length=100)
    y=range(-50, stop=50, length=100)
    g(x, y) = (dot(a, b) - dot(a[1: 2], [x; y]))/a[3]
    plot!(x, y, g,
        label="$(a)'(x-$(b))=0",
        xlabel="x_1", ylabel="x_2", zlabel="x_3",
        st=:surface, camera=camera)

    # # plot the normal vector, since b is on the plane plot the vector b to b+a
    a_x1 = range(b[1], stop=(b[1]+15*a[1]), length=100)
    a_x2 = range(b[2], stop=(b[2]+15*a[2]), length=100)
    a_x3 = range(b[3], stop=(b[3]+15*a[3]), length=100)
    plot!(a_x1, a_x2, a_x3, label="a")
end

function draw_hyperplane_3D(a, b, p, camera=(60, 45))
    pyplot()
    x=range(-2, stop=2, length=100)
    y=range(-2, stop=2, length=100)
    g(x, y) = (dot(a, b) - dot(a[1: 2], [x; y]))/a[3]
    p= plot(x, y, g,
        title="Hyperplane in 3D",
        label="$(a)'(x-$(b))=0",
        xlabel="x_1", ylabel="x_2", zlabel="x_3",
        xlims= (-2, 2), ylims=(-2, 2), zlims=(-2, 2),
        st=:surface, camera=camera)

    # plot the normal vector, since b is on the plane plot the vector b to b+a
    a_x1 = range(b[1], stop=(b[1]+a[1]), length=100)
    a_x2 = range(b[2], stop=(b[2]+a[2]), length=100)
    a_x3 = range(b[3], stop=(b[3]+a[3]), length=100)
    plot!(a_x1, a_x2, a_x3, label="a")
end
```

Out[7]: draw_hyperplane_3D (generic function with 2 methods)

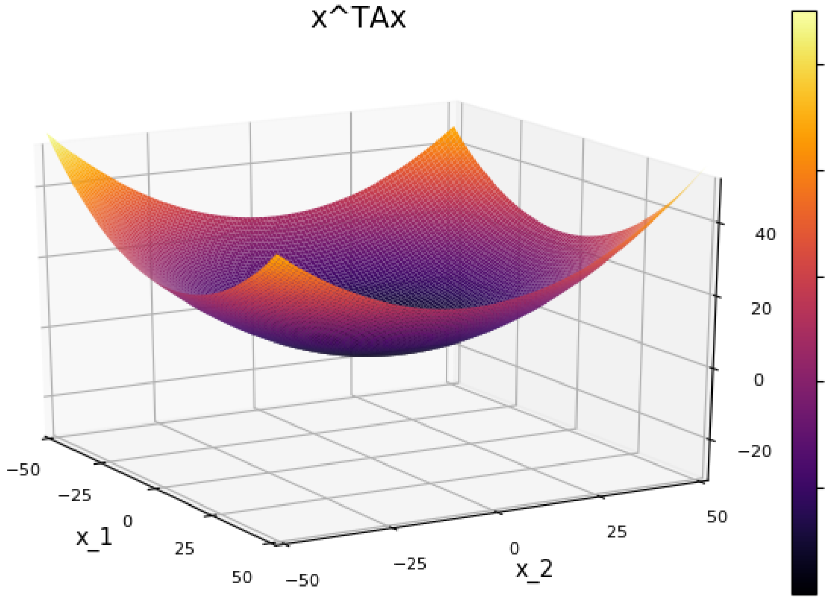
In [8]:
```
function affine_approx(x, f, f_der, x_0)
    return f(x_0) + dot(f_der(x_0), (x-x_0))
end

function f(x)
    return x' * A * x
end

function nabla_f(x)
    return (A+A')x
end
```

Out[8]: nabla_f (generic function with 1 method)

In [11]:

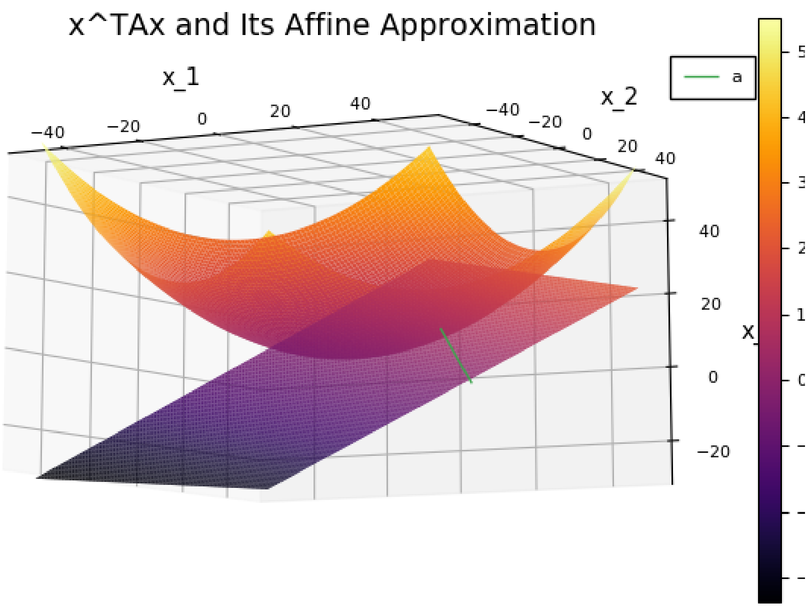
```
camera=(60, 20)
# draw_hyperplane_3D()
pyplot()
x=range(-50, stop=50, length=100)
y=range(-50, stop=50, length=100)
g(x, y) = ([x y] * A * [x; y])[1]
p= plot(x, y, g,
    title="x^TAx",
    label="x^TAx",
    xlabel="x_1", ylabel="x_2",
    zlims=(-30,50),
    st=:surface, camera=camera)
```

Out[11]:



In [12]:
```
camera=(30, -10)


pyplot()
x=range(-50, stop=50, length=100)
y=range(-50, stop=50, length=100)
g(x, y) = ([x y] * A * [x; y])[1]
p= plot(x, y, g,
    title="x^TAx and Its Affine Approximation",
    label="x^TAx",
    xlabel="x_1", ylabel="x_2",
    zlims=(-30,50),
    st=:surface, camera=camera)

x_0 = [25; 0]
a = [nabla_f(x_0); -1]
b = [x_0; f(x_0)]

draw_hyperplane_3D_modified(a, b, camera)
```

Out[12]:



In [ ]: