ECE 8843-A

HOMEWORK-1

Submitted by:

Name: ANIRUDHA SUNDARESAN

GT ID: 903301401

Date: September 1, 2017

Scanned by CamScanner

9① - piazza registered,

Machines are omnipresent and continue to have an increasing role in our decision making processes mainly because they can analyse a lot of data better and more effectively than we can. Everything that we will be doing in the future, will depend on machines. Self-driving cars which utilise the concepts of machine learning with artificial intelligence (AI) are already on the market now. In the future, everything around us will be automated and hence we will come to a point where we will have to rely on machines for basic tasks (Haven't we already!). The main questions that machine learning deals with are these: Given a set of input points and their corresponding output points, can the machine correctly guess the output for an input point not specified in the initial set of input points? Also, if only a set of data points are given, will the machine be able to correctly guess the pattern governing these data points?

What we are learning right now in this course is the mathematical foundations required for machine learning. Machine learning algorithms help the machine learn the input data in an effective way to make decisions. These type of algorithms usually use the probabilistic modelling and the geometric modelling. Both models provide a concise way of representing data points. We need to cover basic mathematical concepts centred on probability, statistics and linear algebra to work with these models. In the first week, we have begun to address the main questions by analysing how a function can be fit to a set of data points. Many problems can then be replaced by this function for further calculations and data modelling. Polynomial curve fitting comes in handy here and we can use the Lagrange interpolation (although limited to lower orders) and Spline polynomial basis functions for this. Using Taylor's expansion formula, any function (originally, not a polynomial) can be expressed as the sum of polynomials. Fourier series converts any random non periodic function to the sum of periodic/ harmonic functions. A polynomial function is easily computable using basic mathematical operators. It is infinitely smooth and also easy to visualize graphically. Hence, polynomial representation of functions is highly preferred. ML will come to play a huge role in a variety of applications.

What makes machine learning so ubiquitous is that it can be used with many other existing branches. For example, in the case of image analysis and pattern recognition, ML algorithms can be used to develop a system to classify unknown images. Such a system is highly useful in making diagnoses and prognoses in the field of healthcare. (http://web.mit.edu/profit/PDFS/EdwardTolson.pdf). Machine learning to be used in the field of trading is highly popular nowadays. Combining ML algorithms with High frequency trading algorithms will effectively decide the number of stocks to buy and also at what price it will yield highest profits. Decision making processes will heavily depend on ML in the years to come. Combining this with AI, a machine can learn to do something of its own accord, the only input needed are some initial state conditions.

Scanned by CamScanner

```
function ft1 = piecepoly2(t,alpha)
% gets an input time vector t.
% gets an input alpha vector alpha.

% The following lines of code are written:
% alpha = [-1 3 2 -1 4] or any alpha you want.
% t = linspace (-5,7,10000) or any time vector you want. The third
% parameter in linspace command gives the sampling rate/ number of samples.
% Higher the value of the rate, smoother is the curve.
% ft = piecepoly2(t,alpha) to get samples of ft as per the question.
% plot(t,ft); xlabel('t'); ylabel('ft') for getting the smooth curve.

for i=1:length(t)
    ft1(i) = alpha(1)*funcb2(t(i)) + alpha(2)*funcb2(t(i)-1) + alpha(3)*funcb2(t(i)-2)
+ alpha(4)*funcb2(t(i)-3) + alpha(5)*funcb2(t(i)-4);
end
end
```


Scanned by CamScanner

```
function b2 = funcb2(t)

% This function calculates the b2(t) according to the question based on the
% values of t.

if (t >=-1.5) && (t<=-0.5)
    b2 = ((t + 1.5)^2)/2;
 else if (t>=-0.5) && (t<=0.5)
        b2 = 0.75 - (t^2);
    else if (t>=0.5) && (t<=1.5)
            b2 = ((t- 1.5)^2)/2;
        else
            b2 = 0;
        end
    end
end
end
```

called by piecepoly2.m

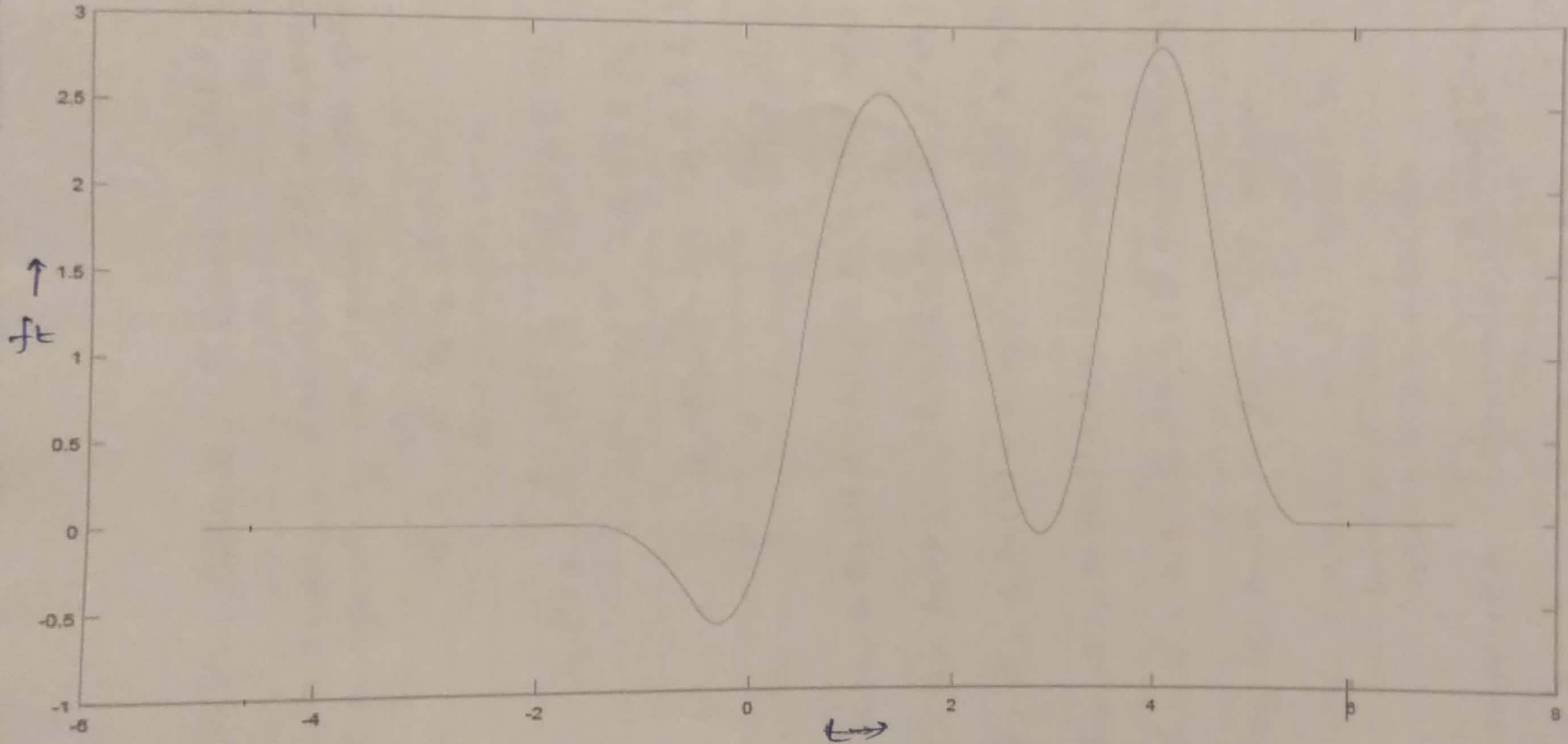
This code calculates:

$$b_2(t) = \begin{cases} (t+3/2)^2/2 & \frac{-3}{2} \le t \le -1/2 \\ -t^2 + 3/4 & -1/2 \le t \le 1/2 \\ (t-3/2)^2/2 & 1/2 \le t \le 3/2 \\ 0 & |t| \ge 3/2 \end{cases}$$

Scanned by CamScanner

3. (a) graph of ft = piecepoly (t, alpha) with high sampling rate

→ for the specific case : $\{-1, 3, 2, -1, 4\}$



Scanned by CamScanner

3(b) $f(0) = 1$ $f(1) = 1/2$ $f(2) = -2$ $f(3) = -3$ $f(4) = -1$

From the question, we know that $f(t)$ is a second order spline that is defined by the overlap of 5B-splines.

$$f(t) = \sum_{k=0}^{4} a_k b_2(t-k)$$

$$b_2(t) = \begin{cases} (t+3/2)^2/2 & -3/2 \le t \le -1/2 \\ -t^2+3/4 & -1/2 \le t \le 1/2 \\ (t-3/2)^2/2 & 1/2 \le t \le 3/2 \\ 0 & |t| \ge 3/2 \end{cases}$$

$f(0) = 1 = a_0 b_2(0) + a_1 b_2(-1) + a_2 b_2(-2) + a_3 b_2(-3) + a_4 b_2(-4)$

$f(1) = 1/2 = a_0 b_2(1) + a_1 b_2(0) + a_2 b_2(-1) + a_3 b_2(-2) + a_4 b_2(-3)$

$f(2) = -2 = a_0 b_2(2) + a_1 b_2(1) + a_2 b_2(0) + a_3 b_2(-1) + a_4 b_2(-2)$

$f(3) = -3 = a_0 b_2(3) + a_1 b_2(2) + a_2 b_2(1) + a_3 b_2(0) + a_4 b_2(-1)$

$f(4) = -1 = a_0 b_2(4) + a_1 b_2(3) + a_2 b_2(2) + a_3 b_2(1) + a_4 b_2(0)$

Using the func b_2. fn from 2(a), we get:

$b_2(-4) = b_2(-3) = b_2(-2) = b_2(2) = b_2(3) = b_2(4) = 0$

$b_2(0) = 0.7500$

$b_2(1) = 0.125 = b_2(-1)$

On substitution: [f matrix] = [B_2 matrix] [alpha]

5×1 5×5 5×1

$$\begin{bmatrix} 1 \\ 0.5 \\ -2 \\ -3 \\ -1 \end{bmatrix} = \begin{bmatrix} 0.75 & 0.125 & 0 & 0 & 0 \\ 0.125 & 0.75 & 0.125 & 0 & 0 \\ 0 & 0.125 & 0.75 & 0.125 & 0 \\ 0 & 0 & 0.125 & 0.75 & 0.125 \\ 0 & 0 & 0 & 0.125 & 0.75 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}$$

Scanned by CamScanner

Taking inverse of B_2 matrix and multiplying with the f matrix yields the alpha matrix.

$$B_2^{-1} \cdot f = Alpha$$

$\therefore$ We get
$a_0 = 1.1937$
$a_1 = 0.8381$
$a_2 = -2.222$
$a_3 = -3.5048$
$a_4 = -0.7492$

3. (c) Generalising 2(b).

$$f(t) = \sum_{n=0}^{N-1} a_n b_2(t-n)$$

Again,

$$f(0) = a_0 b_2(0) + a_1 b_2(-1) + a_2 b_2(-2) + \dots + a_{N-1} b_2(-(N-1))$$

$$f(1) = a_0 b_2(1) + a_1 b_2(0) + a_2 b_2(-1) + \dots + a_{N-1} b_2(2-N)$$

$$\vdots$$

$$f(N-2) = a_0 b_2(N-2) + a_1 b_2(N-3) + \dots + a_{N-1} b_2(-1)$$

$$f(N-1) = a_0 b_2(N-1) + a_1 b_2(N-2) + \dots + a_{N-1} b_2(0)$$

Now, $\underset{N \times 1}{\text{fmatrix}} = [B_2 \text{ matrix}] \; [\text{alpha Matrix}]$

B_2 matrix: $N \times N$ matrix $\approx$ A matrix; alpha Matrix: $N \times 1$

$$\underbrace{\begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(N-2) \\ f(N-1) \end{bmatrix}}_{\text{fmatrix}} = \underbrace{\begin{bmatrix} b_2(0) & b_2(-1) & \cdots & b_2(-(N-1)) \\ b_2(1) & b_2(0) & \cdots & b_2(-(N-2)) \\ \vdots & \vdots & & \vdots \\ b_2(N-2) & b_2(N-3) & \cdots & b_2(1) \\ b_2(N-1) & b_2(N-2) & \cdots & b_2(0) \end{bmatrix}}_{\text{A matrix}} \underbrace{\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{N-2} \\ a_{N-1} \end{bmatrix}}_{\text{alpha matrix}}$$

Scanned by CamScanner

$$\therefore \text{Amatrix} = \begin{bmatrix} b_2(0) & b_2(-1) & \cdots & b_2(-(N-1)) \\ b_2(1) & b_2(0) & & b_2(-(N-2)) \\ \vdots & \vdots & & \vdots \\ b_2(N-2) & b_2(N-3) & \cdots & b_2(-1) \\ b_2(N-1) & b_2(N-2) & \cdots & b_2(0) \end{bmatrix}$$

3.
(d) Suppose that $f(t) = \sum_{n=-\infty}^{\infty} a_n b_2(t-n)$ — ①

$\{a_n\}_{n \in \mathbb{Z}} \rightarrow$ Possibly Infinite sequence of numbers

Replace t by n for more clarification

So, equation ① : $f(n) = \sum_{i=-\infty}^{\infty} a_i b_2(n-i)$ — ②

Given : $f(n) = \sum_{l=-\infty}^{\infty} h_l a_{n-l}$ Let $l = n-k \rightarrow k = n-l$

Changing limits from l to k.

$$\therefore f(n) = \sum_{k=\infty}^{-\infty} h_{n-k} a_k = \sum_{k=-\infty}^{\infty} h_{n+k} a_k \quad — ③$$

$\hookrightarrow (k)$ with $(-k)$

Expanding ②, $f(n) = \cdots + a_0 b_2(n) + a_1 b_2(n-1) + \cdots$
$a_{-1} b_2(n+1)$

Expanding ③ $f(n) = \cdots + h_0(a_n) + h_1 a_{n-1} + \cdots$
$h_n(a_0) + h_{n-1} a_1$

On comparing coefficients:
$\therefore h_n = b_2(n)$; $h_{n-1} = b_2(n-1)$ and so on

Hence
Sequence of numbers $\{h_n\}_{n \in \mathbb{Z}} = \{b_2(n)\}_{n \in \mathbb{Z}}$ $b_2(t)$ as defined in Q. 3.

Scanned by CamScanner

4.
(a) For doing this, I used the polyfit & polyval combo MATLAB.

[nimthorder.m]

→ We get the polynomial as:

$0.001x^9 - 0.0037x^8 + 0.0427x^7 - 0.1616x^6 - 0.7653x^5 + 9.8373x^4 - 38.9988x^3 + 71.9967x^2 - 57.8076x + 13.6239$

→ I also used Lagrangian interpolator (a function that I wrote in MATLAB).

[ninthorderlag.m]

The graphs obtained in both the cases are similar and pass through all the points.

4(b) Uses funcbz.m, ~~piecpoly2(t, alpha)~~ and bspliner.m.

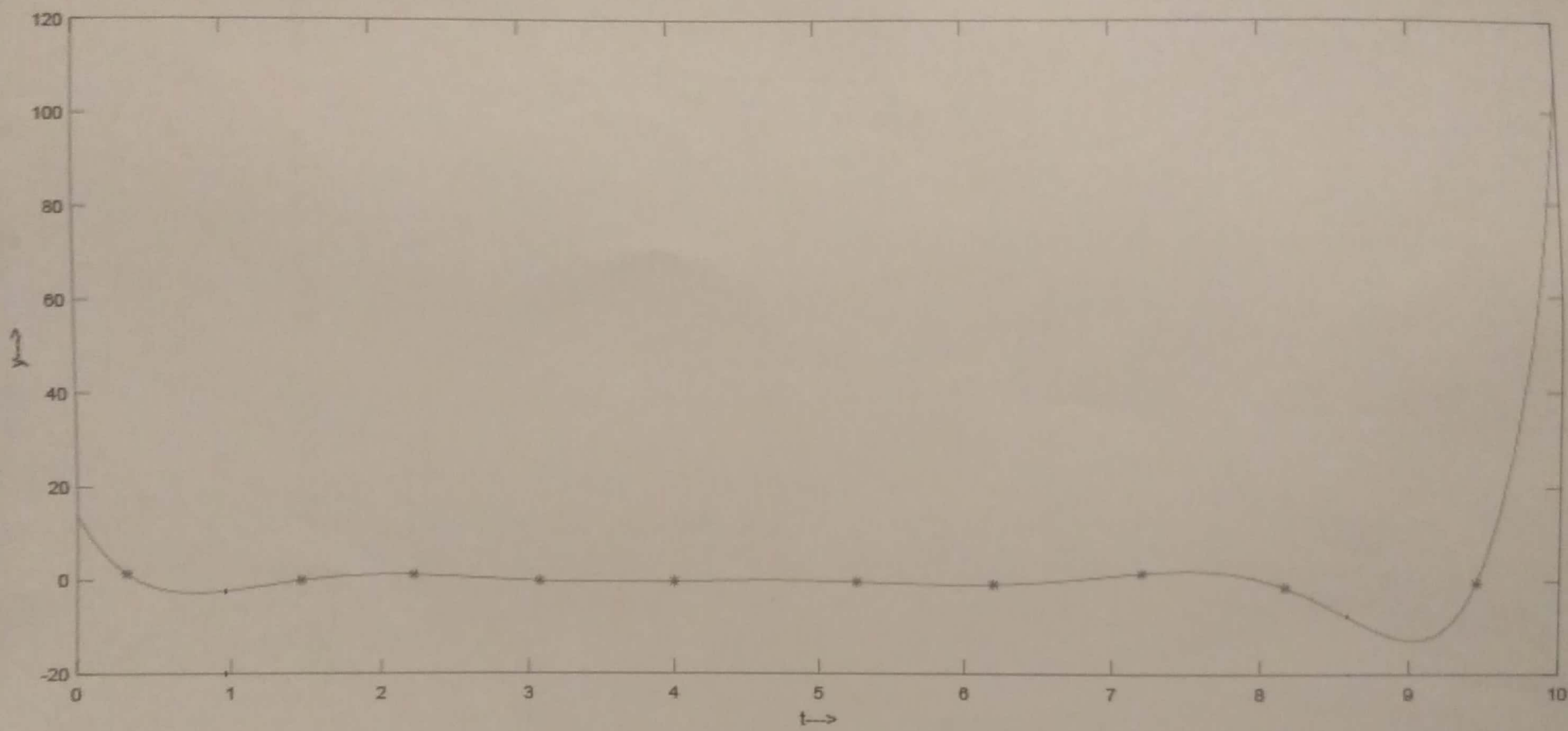
Please see codes and graphs in the following pages.

Scanned by CamScanner

4 (a) Using Lagrangian algorithm.

```
%Finally it should be linspace T and plot T vs f(t) and a(m+1) vector is the y vector.
T = linspace (0,10,100);
x=1;
num =1;
den =1;


for i=1:length(T)
    f(i) = 0;

for m=0:9
    for k=0:9
        if (m~=k)
         b(k+1) = (t(m+1)-t(k+1));
         x(k+1) = T(i) - t(k+1);
        else b(k+1)=1;
            x(k+1)=1;
        end
    end
    den = cumprod(b);
    num = cumprod(x);
    p(m+1) = num(10)/den(10);
end

 for m=0:9
    f(i) = f(i) + a(m+1)*p(m+1);
 end

end

plot(T,f);
xlabel('t--->');
ylabel('y--->');
hold on
plot(t,y,'*');
hold off
```


Scanned by CamScanner



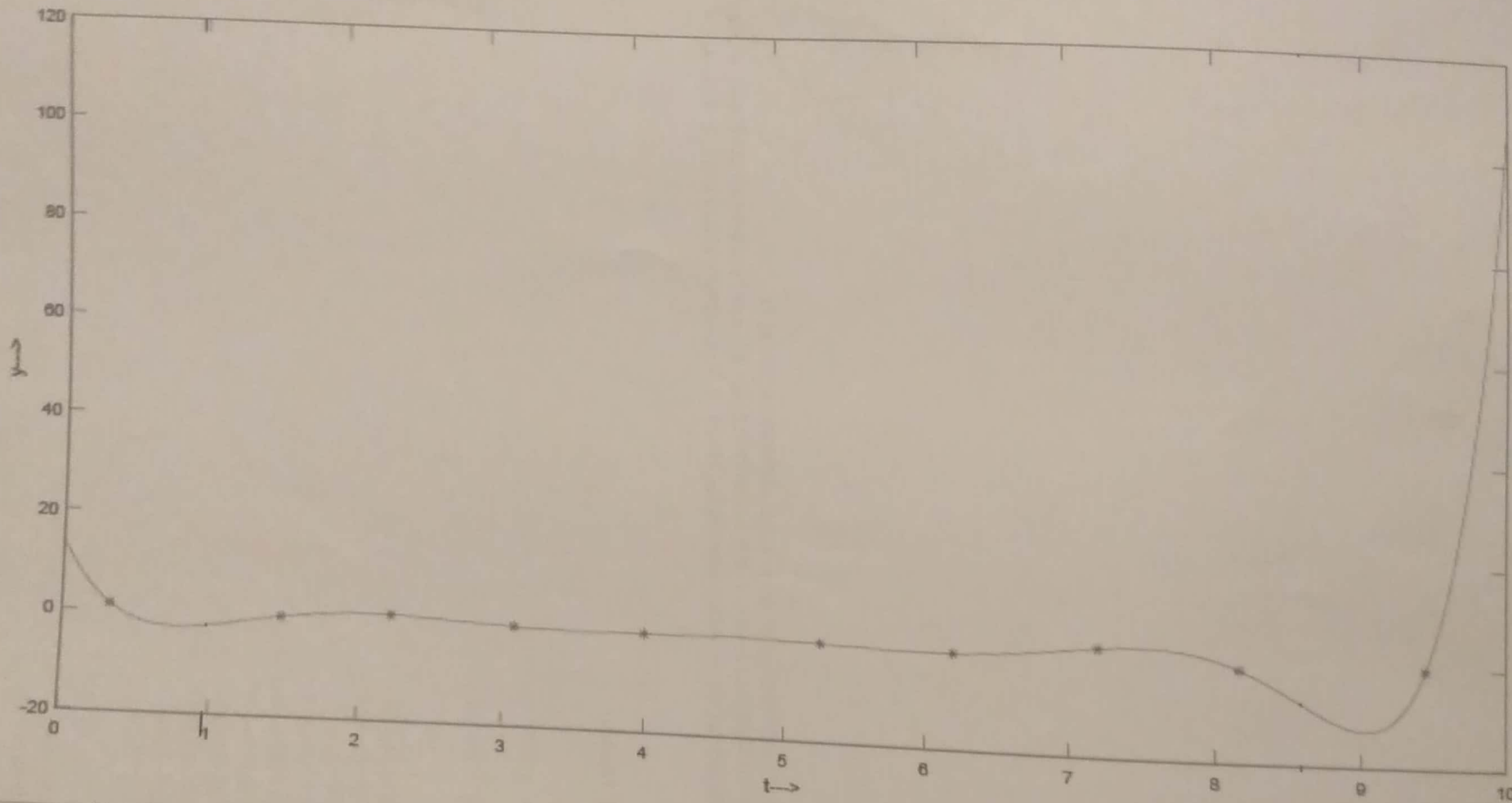
4. (a) Plot using Lagrangian interpolation.
*- sample points overlaid

Scanned by CamScanner

Using inbuilt poly fit & poly val function.

```
coeff = polyfit(t,y,9)
xn = linspace(0,10,1000);
yn = polyval(coeff,xn);
plot(xn,yn)
xlabel('t--->');
ylabel('y--->');
hold on
plot(t,y,'*')
hold off

% coeff = 0.0001   -0.0037    0.0427   -0.1616   -0.7653    9.8373  -38.9788   71.9967
-57.8076   13.6239
```


Scanned by CamScanner



Plot using polyfit + polyval function in MATLAB.

* – sample points

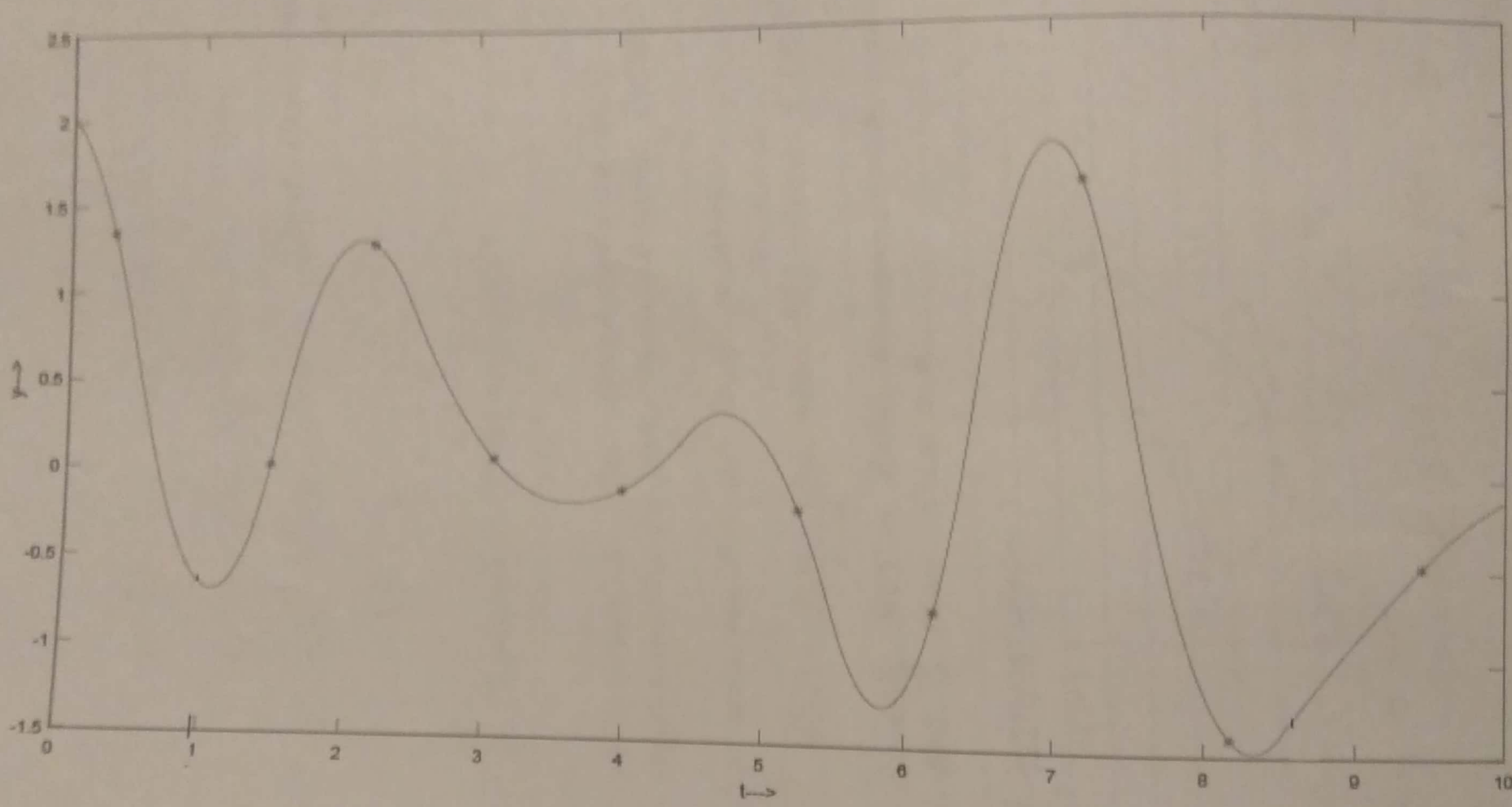
Scanned by CamScanner

```
function ft = bsplinex(t,y)
      d = zeros(10,10);

for i=1:length(t)
    d(i,1) = funcb2(t(i));
    d(i,2) = funcb2(t(i)-1);
    d(i,3) = funcb2(t(i)-2);
    d(i,4) = funcb2(t(i)-3);
    d(i,5) = funcb2(t(i)-4);
    d(i,6) = funcb2(t(i)-5);
    d(i,7) = funcb2(t(i)-6);
    d(i,8) = funcb2(t(i)-7);
    d(i,9) = funcb2(t(i)-8);
    d(i,10) = funcb2(t(i)-9);

    D = d^(-1);
    % Inverse of the A matrix.
    a = D*y;
    % calculting the alpha vector.

end
  T = linspace (0,10,10000);

  for j=1:length(T)

     ft(j) = [a(1)*funcb2(T(j)) + a(2)*funcb2(T(j)-1) + a(3)*funcb2(T(j)-2) + a(4)
*funcb2(T(j)-3) + a(5)*funcb2(T(j)-4) + a(6)*funcb2(T(j)-5) + a(7)*funcb2(T(j)-6) + a
(8)*funcb2(T(j)-7)+ a(9)*funcb2(T(j)-8)+ a(10)*funcb2(T(j)-9)];
  end

 plot(T,ft);
 xlabel('t--->');
 ylabel('y--->');
 hold on
 plot(t,y,'*');

end
```

For 4(b). called by function bsplinex.m

```
function b2 = funcb2(t)

% This function calculates the b2(t) according to the question based on the
% values of t.

if (t >=-1.5) && (t<=-0.5)
    b2 = ((t + 1.5)^2)/2;
  else if (t>=-0.5) && (t<=0.5)
        b2 = 0.75 - (t^2);
    else if (t>=0.5) && (t<=1.5)
            b2 = ((t- 1.5)^2)/2;
        else
              b2 = 0;
        end
    end
end
end
```


Scanned by CamScanner



4 (b) Plot using B-splines

* – sample points

Scanned by CamScanner

5.

(a) Let $f(x) = |x|^p$ for $p \geq 1$. Prove that

$$f\left(\frac{a+b}{2}\right) \leq \frac{f(a)+f(b)}{2}$$

That is, prove that $f(x)$ is convex. (Hint: Mean value theorem)

Answer:

Statement of Mean Value theorem (MVT):

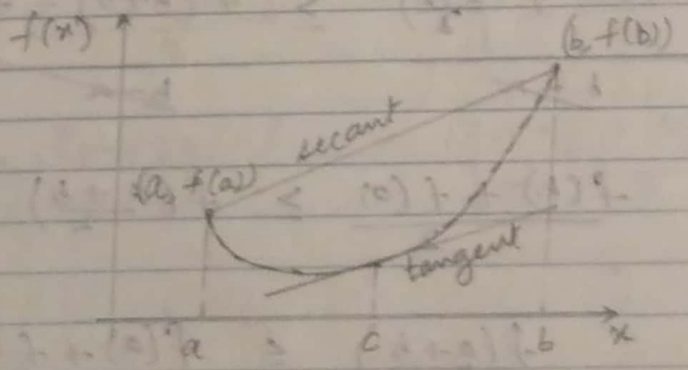
Assume that:
1. f is continuous on the closed interval $[a, b]$
2. f is differentiable on the open interval (a, b)

Then, there is some point c in (a,b) so that

$$f'(c) = \frac{f(b)-f(a)}{b-a} \Rightarrow f(b) - f(a) = f'(c)(b-a)$$

A special case of MVT is Rolle's theorem, which states that if $f(a) = f(b)$, then $f'(c) = 0$

Geometrical interpretation:



So, according to MVT,

secant slope $f'(c) = \dfrac{f(b)-f(a)}{b-a}$

is equal to tangent slope $f'(c)$.

Scanned by CamScanner

Since $f(x) = |x|^p$ for $p \geq 1$

$f'(x) = p \cdot |x|^{p-1} \cdot \frac{x}{|x|} = p \cdot |x|^{p-2} \cdot x$ [$\because$ If $p=1$, $f'(b) = f'(a)$]

If $b > a$, then 3 cases

- $a, b > 0$ — $f'(b) > f'(a)$
- $a < 0$ & $b = 0$ — $f'(b) = 0$, $f'(a) <$
- $a, b < 0$ — $f'(b) > f'(a)$

$\therefore$ $f'(b) \geq f'(a)$ for all $b > a$

Now, let us assume a point c in between the points a & b

$$a < c < b$$

Then, by MVT and because $f'(b) > f'(a)$

$$\frac{f(b) - f(c)}{b - c} \geq \frac{f(c) - f(a)}{c - a}$$

Since $a < c < b$, c can be $= \underline{\frac{a+b}{2} = c}$

Then
$$\frac{f(b) - f\left(\frac{a+b}{2}\right)}{b - \left(\frac{a+b}{2}\right)} \geq \frac{f\left(\frac{a+b}{2}\right) - f(a)}{\frac{a+b}{2} - a}$$

$$\Rightarrow \frac{f(b) - f\left(\frac{a+b}{2}\right)}{\cancel{b-a}} \geq \frac{f\left(\frac{a+b}{2}\right) - f(a)}{\cancel{b-a}} \quad [b > a]$$

$$\longrightarrow \frac{f(b) + f(a)}{2} \geq f\left(\frac{a+b}{2}\right)$$ Rearranging

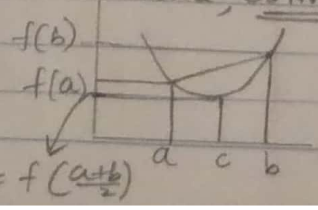
$$\Rightarrow f\left(\frac{a+b}{2}\right) \leq \frac{f(a) + f(b)}{2}$$

Since both $f(a)$ & $f(b)$ greater than $f(c)$, their avg $> f(c)$

Hence, CONVEX.

HENCE PROVED.

This is valid only for a convex function.



5(b) Let S be the set of all (infinite length) sequences

$$S_p = \left\{ \{x_n\}_{n=1}^{\infty} : \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} < \infty \right\}$$

Show that S is indeed a linear vector space.

Answer:

For S to be a linear vector space, it should be closed under scalar multiplication and vector addition.

If $x, y \in S \Rightarrow ax + by \in S$, $\forall a, b \in F$.

then S is a linear vector space over the field F.

Let us take one sample sequence and thus, the condition is:

$$\left(|x_1|^p + |x_2|^p + |x_3|^p + \cdots + |x_\infty|^p \right)^{1/p} < \infty$$

equ ⓐ $\xrightarrow{\text{power } p}$ $|x_1|^p + |x_2|^p + \cdots + |x_\infty|^p < \infty$ $\quad p \geq 1$

Likewise, another sample sequence follows the condition:

$$\left(|y_1|^p + |y_2|^p + |y_3|^p + \cdots + |y_\infty|^p \right)^{1/p} < \infty$$

equation ⓑ $\xrightarrow{\text{power } p}$ $|y_1|^p + |y_2|^p + \cdots + |y_\infty|^p < \infty$ $\quad p \geq 1$

From part(a) of 5th question, $\quad \alpha, \beta$ just scalars

$$\frac{|\alpha x_1|^p + |\beta y_1|^p}{2} \geq \left| \frac{\alpha x_1 + \beta y_1}{2} \right|^p \quad -①$$

Likewise,

$$\frac{|\alpha x_2|^p + |\beta y_2|^p}{2} \geq \left| \frac{\alpha x_2 + \beta y_2}{2} \right|^p \quad -②$$

$$\vdots \qquad \vdots \qquad \vdots$$

and so on,

Finally, $$\frac{|\alpha x_\infty|^p + |\beta y_\infty|^p}{2} \geq \left| \frac{\alpha x_\infty + \beta y_\infty}{2} \right|^p \quad — (\infty)$$

Scanned by CamScanner

Adding equations ①, ②

$$|\alpha x_1|^p + |\alpha x_2|^p + |\alpha x_3|^p + \cdots + |\alpha x_\infty|^p + |\beta y_1|^p + |\beta y_2|^p + \cdots + |\beta y_\infty|^p \geq \frac{1}{2^p}\left[|\alpha x_1 + \beta y_1|^p + |\alpha x_2 + \beta y_2|^p + \cdots + |\alpha x_\infty + \beta y_\infty|^p\right]$$

$$\alpha^p\left[|x_1|^p + |x_2|^p + \cdots |x_\infty|^p\right] + \beta^p\left[|y_1|^p + |y_2|^p + \cdots + |y_\infty|^p\right] \geq \frac{1}{2^p}\left[|\alpha x_1 + \beta y_1|^p + \cdots + |\alpha x_\infty + \beta y_\infty|^p\right]$$

From equation ⓐ and ⓑ L.H.S $< \infty$

$$\therefore |\alpha x_1 + \beta y_1|^p + |\alpha x_2 + \beta y_2|^p + \cdots + |\alpha x_\infty + \beta y_\infty|^p < \infty$$

power $(1/p)$ on both sides to yield

$$\Rightarrow \left[|\alpha x_1 + \beta y_1|^p + |\alpha x_2 + \beta y_2|^p + \cdots + |\alpha x_\infty + \beta y_\infty|^p\right]^{1/p} < \infty$$

$$\left(\sum_{n=1}^{\infty} |\alpha x_n + \beta y_n|^p\right)^{1/p} < \infty$$

Thus, with the initial conditions, ~~with S as vector space~~ we have proved that the main condition satisfies.

Thus, S is indeed a linear vector space.