XIAMEN UNIVERSITY MALAYSIA
厦門大學馬來西亞分校

LINEAR ALGEBRA II PROJECT REPORT

ZENG ZIHANG IBU2209409
WANG FANGYU MAT2309449

October 25, 2024

Contents

Section 1 Question 1: Forest Management

1.1 The model

A harvester has a kind of trees are cut down and replaced with seedlings, maintaining a constant number of trees in the forest. Assuming all seedlings survive and grow, the market values trees based on height, with n price classes. The first class includes seedlings from $[0, h_1)$ with no economic value, and the nth class includes trees taller than or equal to h_{n-1}.

Class	Height Interval	Value (dollars)
1	$[0, h_1)$	$None$
2	$[h_1, h_2)$	P_2
3	$[h_2, h_3)$	P_3
$\vdots$	$\vdots$	$\vdots$
$n-1$	$[h_{n-2}, h_{n-1})$	p_{n-1}
n	$[h_{n-1})$	p_n

Let $x_i, (i = 1, 2, ..., n)$ represent the number of trees in the i-th class that remain after each harvest. We form a column vector with these numbers and call it the nonharvest vector.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

To ensure sustainable harvesting, the forest must be restored to a fixed configuration represented by the nonharvest vector x after each harvst.

The total number of trees,

$$s = x_1 + x_2 + \cdots + x_n \tag{1.1}$$

is determined by available land and tree spacing. After each harvest, the forest is configured by vector x. Trees grow into a new configuration between harvests, and some are removed during the harvest. A seedling replaces each removed tree, restoring the forest to configuration x. Now consider the growth of trees between two harvest periods. We consequently define the following parameters g_i for $i = 1, 2, ..., n-1$:

$$g_i = \text{the fraction of trees in the } i\text{th class that grow into the}$$
$$(i+1)\text{-st class during a growth period}$$
$$1 - g_i = \text{the fraction of trees in the } i\text{th class that remain in the}$$
$$i\text{th class during a growth period}$$

With these $n-1$ growth parameters, we form the following $n \times n$ growth matrixs:

$$\mathbf{G} = \begin{bmatrix} 1-g_1 & 0 & 0 & & \dots & 0 \\ g_1 & 1-g_2 & 0 & & \dots & 0 \\ 0 & g_2 & 1-g_2 & & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1-g_{n-1} & 0 \\ 0 & 0 & 0 & \dots & g_{n-1} & 1 \end{bmatrix} \tag{1.2}$$

Then the numbers of trees in the n classes after the growth period is:

$$\mathbf{Gx} = \begin{bmatrix} (1-g_1)x_1 \\ g_1 + (1-g_2)x_2 \\ g_2x_2 + (1-g_3)x_3 \\ \vdots \\ g_{n-2}x_{n-2} + (1-g_{n-1})x_{n-1} \\ g_{n-1}x_{n-1} + x_n \end{bmatrix} \tag{1.3}$$

Suppose that during the harvest we remove y_i(i=1,2,...,n) trees from the ith class.We get the column vector:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Thus, a total of

$$y_1 + y_2 + \cdots + y_n$$

trees are removed at each harvest. Then the column vector

$$\mathbf{y}' = \begin{bmatrix} y_1 + y_2 + \cdots + y_n \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{1.4}$$

specifies the configuration of trees planted after each harvest.

Now, we have

Gx: The number of trees configuration at end of growth period.

y: The number of trees that harvest.

y': The number of new seedling replacement after harvesting.

x: Configuration at beginning of growth period.

We have the equation below:

$$Gx - y + y' = x$$

$$(I - R)y = (G - I)x \tag{1.5}$$

$$\begin{bmatrix} 0 & -1 & -1 & \dots & -1 & -1 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix} = \begin{bmatrix} -g_1 & 0 & 0 & \dots & 0 & 0 \\ g_1 & -g_2 & 0 & \dots & 0 & 0 \\ 0 & g_2 & -g_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -g_{n-1} & 0 \\ 0 & 0 & 0 & \dots & g_{n-1} & 0 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}$$

Because the value of y_1 doesn't affect the calculation results, so we assume $y_1 = 0$. Then,

$$\begin{aligned} y_2 + y_3 + \cdots + y_n &= g_1 x_1 \\ y_2 &= g_1 x_1 - g_2 x_2 \\ y_3 &= g_2 x_2 - g_3 x_3 \\ &\vdots \\ y_{n-1} &= g_{n-2} x_{n-2} - g_{n-1} x_{n-1} \\ y_n &= g_{n-1} x_{n-1} \end{aligned} \tag{1.6}$$

Because we must have $y_i \geq 0$ for $i = 2, 3, \dots, n$, we require that

$$g_1x_1 \geq g_2x_2 \geq \cdots \geq g_{n-1}x_{n-1} \geq 0 \tag{1.7}$$

1.2 Optimal Sustainable Yield

The total yield of the harvest, Yld, is given by

$$Yld = p_2y_2 + p_3y_3 + \cdots + p_ny_n \tag{1.8}$$

Using(1.6) we may substitute for the y_i's in (1.8) to obtain

$$Yld = p_2g_1x_1 + (p_3 - p_2)\, g_2x_2 + \cdots + (p_n - p_{n-1})\, g_{n-1}x_{n-1} \tag{1.9}$$

Combining (1.9), (1.1), and (1.7), we can now state the problem of maximizing the yield of the forest over all possible sustainable harvesting policies as follows:

Problem 1. Find the nonnetative number $x_1, x_2, ..., x_n$ that maximize.

$$Yld = p_2g_1x_1 + (p_3 - p_2)\, g_2x_2 + \cdots + (p_n - p_{n-1})\, g_{n-1}x_{n-1}$$

subject to

$$x_1 + x_2 + \cdots x_n = s$$

and

$$g_1x_1 \geqslant g_2x_2 \geqslant \cdots \geqslant g_{n-1}x_{n-1} \geqslant 0$$

As formulated above, this problem belongs to the field of linear programming. However, we will illustrate the following result, without linear programming theory, by actually exhibiting a sustainable harvesting policy.

Theorem 1.1. ***Optimal Sustainable Yield***
The optimal sustainable yield is achieved by harvesting all the trees from one particular height class and none of the trees from any other height class.

Let us first set

$$\begin{aligned} Yld_k &= yield\ obtained\ by\ harvesting\ all\ of \\ &\quad the\ kth\ calss\ and\ none\ of\ other\ classes \end{aligned}$$

The largest value of Yld_k for $k = 2, 3, ..., n$ will then be the optimal sustainable yield, and the corresponding value of k will be the class that should be completely harvested to attain the optimal sustainable yield.Because no class but the kth is harvested, we have

$$y_2 = y_3 = \cdots = y_{k-1} = y_{k+1} = \cdots = y_n = 0 \tag{1.10}$$

In addition, because all of the kth class is harvested, no trees are present in the height classes above the kth class. Thus,

$$x_k = x_{k+1} = \cdots = x_n = 0 \tag{1.11}$$

Substituting (1.10) and (1.11) into the sustainable harvesting condition (1.8) gives

$$\begin{aligned} y_k &= g_1 x_1 \\ 0 &= g_1 x_1 - g_2 x_2 \\ 0 &= g_2 x_2 - g_3 x_3 \\ &\vdots \\ 0 &= g_{k-2} x_{k-2} - g_{k-1} x_{k-1} \\ y_k &= g_{k-1} x_{k-1} \end{aligned} \tag{12}$$

Equation(1.12) can also be written as

$$y_k = g_1 x_1 = g_2 x_2 = \cdots = g_{k-1} x_{k-1} \tag{1.13}$$

from which is follows that

$$\begin{aligned} x_2 &= g_1 x_1 / g_2 \\ x_3 &= g_1 x_1 / g_3 \\ &\vdots \\ x_{k-1} &= g_1 x_1 / g_{k-1} \end{aligned} \tag{1.14}$$

If we substitute Equation (1.11) and (1.14) into

$$x_1 + x_2 + \cdots + x_n = s$$

[which is Equation(1.1)], we can solve for x_1 and obtain

$$x_1 = \frac{s}{1 + \frac{g_1}{g_2} + \frac{g_1}{g_3} + \cdots + \frac{g_1}{g_{k-1}}} \tag{1.15}$$

For the yield Yld_k, we combine(1.8),(1.10),(1.13),and (1.15) to obtain

$$\begin{aligned} Yld_k &= p_2y_2 + p_3y_3 + \cdots + p_ny_n \\ &= p_ky_k \\ &= p_kg_1x_1 \\ &= \frac{p_ks}{\frac{1}{g_1} + \frac{1}{g_2} + \cdots + \frac{1}{g_{k-1}}} \end{aligned} \tag{1.16}$$

Equation(1.16) determines Yld_k in terms of the known growth and economic parameters for any $k = 2, 3, ..., n$. Thus, the optimal sustainable yield is found as follows.

Theorem 1.2. ***Finding the Optimal Sustainable Yield***
The optimal sustainable yield is the largest value of

$$\frac{p_ks}{\frac{1}{g_1} + \frac{1}{g_2} + \cdots + \frac{1}{g_{k-1}}}$$

for $k = 2, 3, ..., n$. The corresponding value of k is the number of the class that is completely harvested.

Theorem (1.2) implies that it is not necessarily the highest-priced class of trees that should be totally cropped. The growth parameters g_i must also be taken into account to determine the optimal sustainable yield.

1.3 Problem

A forestry company cultivates an economically valuable tree species. The growth parameters for these trees are as follows:

$$g_i = \frac{1}{e^i}$$

for i = 1,2,3,...,n-1, where n=10 (n is the total number of height classes) can be chosen as large as needed. The total number of trees is a constant s.Suppose that the value and cost (in dollars) of a tree in the kth height interval are respectively given by

$$v_k = 0.002\rho^k,$$
$$c_k = 0.003e^{2k}$$

where ρ is a parameter satisfying $1 \leq \rho \leq 9$, and k takes only integer values.
Determine the value of ρ and k for which the company is profitable and when the profit is maximized.

Solution The profit of a tree is

$$p_k = v_k - c_k = 0.002\rho^k - 0.003e^{2k}$$

From equation (1.16) we have,

$$Yld_k = \frac{p_k s}{g_1^{-1} + g_2^{-1} + \cdots + g_{k-1}^{-1}}$$

Now we substitute the given information in the R.H.S. of Yld_k to obtain

$$Yld_k = \frac{(0.002\rho^k - 0.003e^{2k})s}{e^1 + e^2 + e^3 + \cdots + e^{k-1}}$$

As the denominator part of the above expression is in G.P. with common ratio e.
Therefore,

$$e^1 + e^2 + e^3 + \cdots + e^{k-1} = \frac{e(e^{k-1} - 1)}{e - 1}$$

Then we get

$$Yld_k = \frac{(e-1)(0.002\rho^k - 0.003e^{2k})s}{e^k - e}$$

Since s is constant, so the expression of Yld_k depends only on k and ρ . We use MATLAB to run the integer value k from 2 to 9 and try to observe for which value of ρ, Yld_k is maximum for a certain value of k. By using the MATLAB code

```
k_values = [2, 3, 4, 5, 6, 7, 8, 9];
rho_values = [1, 2, 3, 4, 5, 6, 7, 8, 9];

result_matrix = zeros(length(x_values), length(k_values));
    for i = 1:length(x_values)
        rho = rho_values(i);
        for j = 1:length(k_values)
            k = k_values(j);
            result_matrix(i, j) = (exp(1)-1) *
            (0.002*rho^k - 0.003*exp(2*k)) / (exp(k) - exp(1));
        end
    end

disp(result_matrix);
```

we can easily observe that the company is profitable under the following conditions. For ρ = 8, Yld_k is positive only for $k \geq 5$. For ρ = 9, each Yld_k is positive for a certain value of k except k = 1. When $\rho = 9$ and $k = 9$, the company's profit is maximized, observed to be 122.58s dollars.

Section 2 Question 2: Cubic Spline

2.1 Introduction

The estimated mean atmospheric concentration of carbon dioxide in Earth's atmosphere is given in the following table:

Since we have seven pairs of data, label these seven pairs of data as seven points such as

$$\left(x_1, y_1\right), \left(x_2, y_2\right), ..., \left(x_n, y_n\right)$$

x_1 to x_7 be 1850 to 2000, y_i be the corresponding concentration carbon dioxide of x_i.

So we have

$$x_2 - x_1 = x_3 - x_2 = \cdots = x_7 - x_6 = h = 25$$

Let $y = S(x)$, by the linear beamer theory, we have

$$S^{(iv)}\left(x\right) \equiv 0 \tag{2.1}$$

$S\left(x\right), S'\left(x\right), S''\left(x\right)$ are continuous for $x_1 \leqslant x \leqslant x_7$.

Devide $[x_1, x_7]$ into 6 equal sub-intervals

$$\left[x_1, x_2\right], \left[x_2, x_3\right], ..., \left[x_6, x_7\right]$$

Since $S^{(iv)}\left(x\right) \equiv 0$, so $S\left(x\right)$ must be a cubic polynomial in x in each sub-interval, Thus

$$S\left(x\right) = \begin{cases} S_1\left(x\right), x_1 \leqslant x \leqslant x_2 \\ S_2\left(x\right), x_2 \leqslant x \leqslant x_3 \\ \vdots \\ S_6\left(x\right), x_6 \leqslant x \leqslant x_7 \end{cases} \tag{2.2}$$

$S_1\left(x\right), S_2\left(x\right), ..., S_6\left(x\right)$ are cubic polynomials. For convenience

$$S_1\left(x\right), S_2\left(x\right), ..., S_6\left(x\right)$$

$$\begin{aligned} S_1\left(x\right) &= a_1\left(x - x_1\right)^3 + b_1\left(x - x_1\right)^2 + c_1\left(x - x_1\right) + d_1, x_1 \leqslant x \leqslant x_2 \\ S_2\left(x\right) &= a_2\left(x - x_2\right)^3 + b_2\left(x - x_2\right)^2 + c_2\left(x - x_2\right) + d_2, x_2 \leqslant x \leqslant x_3 \\ &\vdots \\ S_6\left(x\right) &= a_6\left(x - x_6\right)^3 + b_6\left(x - x_6\right)^2 + c_6\left(x - x_6\right) + d_6, x_6 \leqslant x \leqslant x_7 \end{aligned} \tag{2.3}$$

Since $S'(x), S''(x)$ are continuous, Thus

$$\begin{aligned}
S'_1(x) &= 3a_1(x-x_1)^2 + 2b_1(x-x_1) + c_1, x_1 \leqslant x \leqslant x_2 \\
S'_2(x) &= 3a_2(x-x_2)^2 + 2b_2(x-x_2) + c_2, x_2 \leqslant x \leqslant x_3 \\
&\vdots \\
S'_6(x) &= 3a_6(x-x_6)^2 + 2b_6(x-x_6) + c_6, x_6 \leqslant x \leqslant x_7
\end{aligned} \tag{2.4}$$

and

$$\begin{aligned}
S''_1(x) &= 6a_1(x-x_1) + 2b_1, x_1 \leqslant x \leqslant x_2 \\
S''_2(x) &= 6a_2(x-x_2) + 2b_2, x_2 \leqslant x \leqslant x_3 \\
&\vdots \\
S''_6(x) &= 6a_6(x-x_6) + 2b_6, x_6 \leqslant x \leqslant x_7
\end{aligned} \tag{2.5}$$

Since

$$S(x_1) = y_1, S(x_2) = y_2, ..., S(x_7) = y_7 \tag{2.6}$$

and from the first $n-1$ of these equation and (2.3), we obtain

$$\begin{aligned}
d_1 &= y_1 \\
d_2 &= y_2 \\
&\vdots \\
d_6 &= y_6
\end{aligned} \tag{2.7}$$

From the last equation in (2.6) and last equation in (2.3), and the fact that $x_7 - x_6 = h$, we obtain

$$a_6h^3 + b_6h^2 + c_6h + d_6 = y_7 \tag{2.8}$$

Since $S(x)$ is continuous, Thus

$$S_{i-1}(x_1) = S_i(x_1) = y_i, i = 2, 3, ..., 6$$

So

$$\begin{aligned} a_1h^3 + b_1h^2 + c_1h + d_1 &= y_2 \\ a_2h^3 + b_2h^2 + c_2h + d_2 &= y_3 \\ &\vdots \\ a_5h^3 + b_5h^2 + c_5h + d_5 &= y_6 \end{aligned} \tag{2.9}$$

$S'(x)$ is continuous,Thus

$$S'_{i-1}(x_i) = S'_i(x_i), i = 2, 3, ..., 6$$

So

$$\begin{aligned} 3a_1h^2 + 2b_1h + c_1 &= c_2 \\ 3a_2h^2 + 2b_2h + c_2 &= c_3 \\ &\vdots \\ 3a_5h^2 + 2b_5h + c_5 &= c_6 \end{aligned} \tag{2.10}$$

$S''(x)$ is continuous, Thus

$$S''_{i-1}(x_i) = S''_i(x_i), i = 2, 3, ..., 6$$

So

$$\begin{aligned} 6a_1h + 2b_1 &= 2b_2 \\ 6a_2h + 2b_2 &= 2b_3 \\ &\vdots \\ 6a_5h + 2b_5 &= 2b_6 \end{aligned} \tag{2.11}$$

Let $M_1 = S''(x_1), M_2 = S''(x_2), ..., M_7 = S''(x_7)$. Since (2.6),Thus

$$\begin{aligned} M_1 &= 2b_1 \\ M_2 &= 2b_2 \\ &\vdots \\ M_6 &= 2b_6 \end{aligned} \tag{2.12}$$

So

$$b_1 = \frac{1}{2}M_1, b_2 = \frac{1}{2}M_2, ..., b_6 = \frac{1}{2}M_6 \tag{2.13}$$

So we have

$$\begin{aligned} a_i &= \left(M_{i+1} - M_i\right)/6h \\ b_i &= M_i/2 \\ c_i &= \left(y_{i+1} - y_i\right)/h - \left[\left(M_{i+1} + 2M_i\right)h/6\right] \\ d_i &= y_i \end{aligned} \tag{2.14}$$

for $i = 1, 2, ..., 6$.

Substitute the expression $a_i, b_i, and c_i$ into (2.11), we obtain

$$\begin{gathered} M_1 + 4M_2 + M_3 = 6\left(y_1 - 2y_2 + y_3\right)/h^2 \\ M_2 + 4M_3 + M_4 = 6\left(y_2 - 2y_3 + y_4\right)/h^2 \\ \vdots \\ M_5 + 4M_6 + M_7 = 6\left(y_5 - 2y_6 + y_7\right)/h^2 \end{gathered} \tag{2.15}$$

or in matrix form

$$\begin{bmatrix} 1 & 4 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \\ M_5 \\ M_6 \\ M_7 \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ y_4 - 2y_5 + y_6 \\ y_5 - 2y_6 + y_7 \end{bmatrix} \tag{2.16}$$

2.2 Natural Spline

Natural spline means the second derivative of the spline is zero at the endpoints, which means

$$M_1 = M_7 = 0$$

So Equation (2.16) can be written as

$$\begin{bmatrix} 4 & 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} M_2 \\ M_3 \\ M_4 \\ M_5 \\ M_6 \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ y_4 - 2y_5 + y_6 \\ y_5 - 2y_6 + y_7 \end{bmatrix}$$

$$M_1 = M_7 = 0$$

So the system has unique solution

$$\begin{bmatrix} M_2 \\ M_3 \\ M_4 \\ M_5 \\ M_6 \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} 4 & 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}^{-1} \begin{bmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ y_4 - 2y_5 + y_6 \\ y_5 - 2y_6 + y_7 \end{bmatrix}$$

Here is the Matlab codes to :

1.Solve the M matrix

2.Solve the coefficient a, b, c, d

3.Estimate the CO2 concentration in 1990 and 2010

4.Plot the relevant natural spline

```
x = [1850, 1875, 1900, 1925, 1950, 1975, 2000];
y = [285.2, 288.6, 295.7, 305.3, 311.3, 331.36, 369.64];
h = 25;
n = length(x) - 1;

z = zeros(1, n-1);
for i = 1:n-1
```

```
    z(i) = (y(i) - 2 * y(i+1) + y(i+2)) * 6 / (h^2);
end

M = zeros(1, n+1);
a = zeros(1, n);
b = zeros(1, n);
c = zeros(1, n);
d = zeros(1, n);

z_matrix = z';

matrix = diag(4*ones(1,n-1)) + diag(ones(1,n-2), 1) + diag(ones(1,n-2),

M_matrix = inv(matrix) * z_matrix;
for i = 1:n-1
    M(i+1) = M_matrix(i);
end
M(1) = 0;
M(n+1) = 0;

disp('M=');
disp(M);

for i = 1:n
    a(i) = (M(i+1) - M(i)) / (6 * h);
    b(i) = M(i) / 2;
    c(i) = (y(i+1) - y(i)) / h - h * (M(i+1) + 2 * M(i)) / 6;
    d(i) = y(i);
end

disp('a=');
disp(a);
```

```
disp('b=');
disp(b);
disp('c=');
disp(c);
disp('d=');
disp(d);

year_estimates = [1990, 2010];
spline_estimates = zeros(1, 2);
for j = 1:2
    year = year_estimates(j);
    i = floor((year - 1850) / h) + 1;
    if i > n
        i = n;
    end
    dx = year - x(i);
    spline_estimates(j) = a(i) * dx^3 + b(i) * dx^2 + c(i) * dx + d(i);
end
disp('CO2 concentration estimates for 1990 and 2010:');
disp(spline_estimates);

xx = linspace(x(1), x(end), 1000);
yy = zeros(1, length(xx));
for k = 1:length(xx)
    year = xx(k);
    i = floor((year - 1850) / h) + 1;
    if i > n
        i = n;
    end
    dx = year - x(i);
    yy(k) = a(i) * dx^3 + b(i) * dx^2 + c(i) * dx + d(i);
end
```

```
figure;
plot(x, y, 'o', xx, yy, '-');
title('Natural Cubic Spline Interpolation of CO2 Concentration');
xlabel('Year');
ylabel('CO2 Concentration (ppm)');
legend('Data points', 'Cubic spline', 'Location', 'Best');
grid on;
```

The output of Estimations of the CO2 concentration in 1990 and 2010 are

CO2 concentration estimates for 1990 and 2010:

353.0528 386.2272

The natural spline is

2.3 The Parabolic Runout Spline

The parabolic runout spline means the spline reducesto a parabolic curveon the 5rst and lastintervals, which means

$$M_1 = M_2 \qquad M_6 = M_7$$

So Equation (2.16) can be written as

$$\begin{bmatrix} 5 & 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix} \begin{bmatrix} M_2 \\ M_3 \\ M_4 \\ M_5 \\ M_6 \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ y_4 - 2y_5 + y_6 \\ y_5 - 2y_6 + y_7 \end{bmatrix}$$

$$M_1 = M_7 = 0$$

So the system has unique solution

$$\begin{bmatrix} M_2 \\ M_3 \\ M_4 \\ M_5 \\ M_6 \end{bmatrix} = \frac{6}{h^2} \begin{bmatrix} 5 & 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 5 \end{bmatrix}^{-1} \begin{bmatrix} y_1 - 2y_2 + y_3 \\ y_2 - 2y_3 + y_4 \\ y_3 - 2y_4 + y_5 \\ y_4 - 2y_5 + y_6 \\ y_5 - 2y_6 + y_7 \end{bmatrix}$$

Here is the Matlab codes to :

1.Solve the M matrix

2.Solve the coefficient a, b, c, d

3.Estimate the CO2 concentration in 1990 and 2010

4.Plot the relevant natural spline

```
x = [1850, 1875, 1900, 1925, 1950, 1975, 2000];
y = [285.2, 288.6, 295.7, 305.3, 311.3, 331.36, 369.64];
h = 25;
n = length(x) - 1;

z = zeros(1, n-1);
```

```
for i = 1:n-1
    z(i) = (y(i) - 2 * y(i+1) + y(i+2)) * 6 / (h^2);
end

M = zeros(1, n+1);
a = zeros(1, n);
b = zeros(1, n);
c = zeros(1, n);
d = zeros(1, n);

z_matrix = z';

matrix = [5 1 0 0 0;
          1 4 1 0 0;
          0 1 4 1 0;
          0 0 1 4 1;
          0 0 0 1 5];

inverse_matrix = inv(matrix);
M_matrix = inverse_matrix * z_matrix;
for i = 1:n-1
    M(i+1) = M_matrix(i);
end
M(1) = M(2);
M(n+1) = M(n);

disp('M=');
disp(M);

for i = 1:n
    a(i) = (M(i+1) - M(i)) / (6 * h);
    b(i) = M(i) / 2;
```

```
        c(i) = (y(i+1) - y(i)) / h - h * (M(i+1) + 2 * M(i)) / 6;
        d(i) = y(i);
    end

    % Display a, b, c, d
    disp('a=');
    disp(a);
    disp('b=');
    disp(b);
    disp('c=');
    disp(c);
    disp('d=');
    disp(d);

    year_estimates = [1990, 2010];
    spline_estimates = zeros(1, 2);
    for j = 1:2
        year = year_estimates(j);
        i = floor((year - 1850) / h) + 1;
        if i > n
            i = n;
        end
        dx = year - x(i);
        spline_estimates(j) = a(i)*dx^3 + b(i)*dx^2 + c(i)*dx + d(i);
    end
    disp('CO2 concentration estimates for 1990 and 2010:');
    disp(spline_estimates);

    xx = linspace(x(1), x(end), 1000);
    yy = zeros(1, length(xx));
    for k = 1:length(xx)
        year = xx(k);
```

```
    i = floor((year - 1850) / h) + 1;
    if i > n
        i = n;
    end
    dx = year - x(i);
    yy(k) = a(i)*dx^3 + b(i)*dx^2 + c(i)*dx + d(i);
end
figure;
plot(x, y, 'o', xx, yy, '-');
title('Parabolic Runout Cubic Spline Interpolation of CO2 Concentration
xlabel('Year');
ylabel('CO2 Concentration (ppm)');
legend('Data points', 'Cubic spline', 'Location', 'Best');
grid on;
```

The output of Estimations of the CO2 concentration in 1990 and 2010 are

CO2 concentration estimates for 1990 and 2010:

352.1733 389.9797

The natural spline is

2.4 Comparison of Parabolic Runout Spline's ,Natural Spline's estimated results to real data

This website provides real data on carbon dioxide concentration provided by NASA: `https://data.giss.nasa.gov/modelforce/ghgases/Fig1A.ext.txt`. And the real data is been shown as follows.

The real data of the CO2 concentration in 1990 and 2010 are

CO2 concentration of 1990 and 2010:

354.29 389.21

The following table shows the comparison between the real data and the estimated data obtained by the Parabolic Runout Spline and Natural Spline.

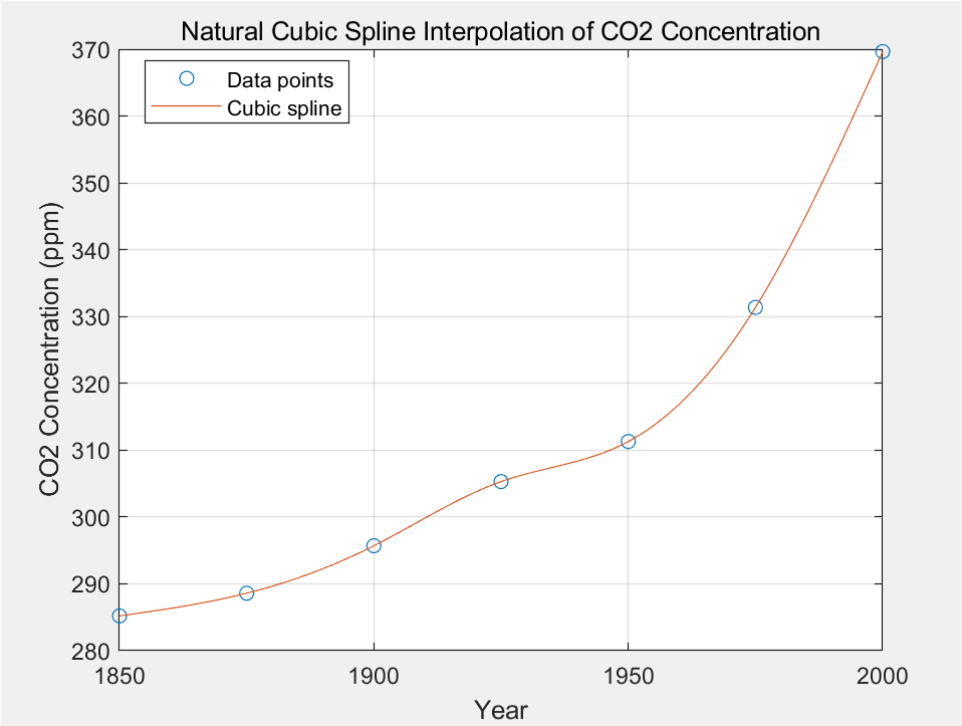
Cubic Splines/Real data	Concentration	
	1990	2010
Natural Spline	353.0528	386.2272
Parabolic Runout Spline	352.1733	389.9797
Real data	354.29	389.21

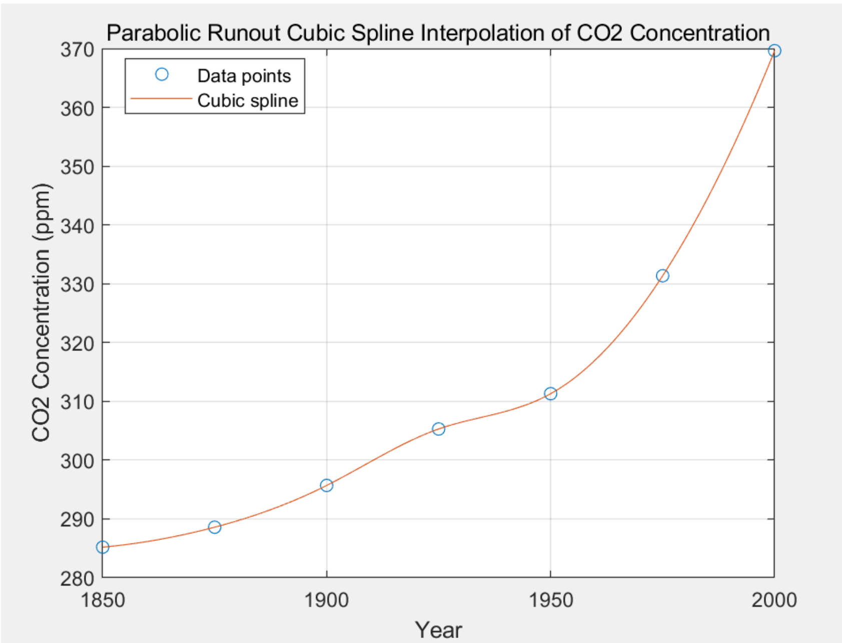
Table 1: Comparison of Carbon Dioxide Concentrations

Through the Table 1, we can visually see that the predictions of the two fitting schemes and the real data are very close. Thus the cubic splines have a significant advantage in data fitting.

333.png

Year	CO2 (ppm)
1850	285.2
1875	288.6
1900	295.7
1925	305.3
1950	311.3
1975	331.36
2000	369.64


Natural Cubic Spline Interpolation of CO2 Concentration
Data points
Cubic spline
CO2 Concentration (ppm)
Year
370
360
350
340
330
320
310
300
290
280
1850
1900
1950
2000


Parabolic Runout Cubic Spline Interpolation of CO2 Concentration
Data points
Cubic spline
CO2 Concentration (ppm)
Year
370
360
350
340
330
320
310
300
290
280
1850
1900
1950
2000

Global Mean CO2 Mixing Ratios (ppm): Observations

Data Source	Year	MixR	Yar	MixR	Data Source	Year	MixR		Year	MixR
Ice-	1850	285.2	1900	295.7		1950	311.3		2000	369.64
Core	1851	285.1	1901	296.2		1951	311.8		2001	371.15
Data	1852	285.0	1902	296.6		1952	312.2		2002	373.15
Adjus-	1853	285.0	1903	297.0		1953	312.6		2003	375.64
ted	1854	284.9	1904	297.5		1954	313.2	NOAA/	2004	377.44
for	1855	285.1	1905	298.0		1955	313.7	ESRL/	2005	379.46
Global	1856	285.4	1906	298.4		1956	314.3	trends	2006	381.59
Mean	1857	285.6	1907	298.8		1957	314.8	change	2007	383.37
	1858	285.9	1908	299.3	SIO	1958	315.34	added	2008	385.46
	1859	286.1	1909	299.7	Mauna	1959	316.18	to	2009	386.95
	1860	286.4	1910	300.1	Loa	1960	317.07	2003	2010	389.21
	1861	286.6	1911	300.6	&	1961	317.73	data	2011	391.15
	1862	286.7	1912	301.0	South	1962	318.43			
	1863	286.8	1913	301.3	Pole	1963	319.08			
	1864	286.9	1914	301.4	Adjus-	1964	319.65			
	1865	287.1	1915	301.6	ted	1965	320.23			
	1866	287.2	1916	302.0	for	1966	321.59			
	1867	287.3	1917	302.4	Global	1967	322.31			
	1868	287.4	1918	302.8	Mean	1968	323.04			
	1869	287.5	1919	303.0		1969	324.23			
	1870	287.7	1920	303.4		1970	325.54			
	1871	287.9	1921	303.7		1971	326.42			
	1872	288.0	1922	304.1		1972	327.45			
	1873	288.2	1923	304.5		1973	329.43			
	1874	288.4	1924	304.9		1974	330.21			
	1875	288.6	1925	305.3	CMDL	1975	331.36			
	1876	288.7	1926	305.8	InSitu	1976	331.92			
	1877	288.9	1927	306.2	Mauna	1977	333.73			
	1878	289.5	1928	306.6	Loa	1978	335.42			
	1879	290.1	1929	307.2	&	1979	337.10			
	1880	290.8	1930	307.5	South	1980	338.99			
	1881	291.4	1931	308.0	Pole	1981	340.36			
	1882	292.0	1932	308.3		1982	341.57			
	1883	292.5	1933	308.9	CMDL	1983	342.53			
	1884	292.9	1934	309.3	Flask	1984	344.24			
	1885	293.3	1935	309.7	Mean	1985	345.72			
	1886	293.8	1936	310.1	of	1986	347.15			
	1887	294.0	1937	310.6	Many	1987	348.93			
	1888	294.1	1938	311.0	Sites	1988	351.47			
	1889	294.2	1939	311.2		1989	353.15			
	1890	294.4	1940	311.3		1990	354.29			
	1891	294.6	1941	311.0		1991	355.68			
	1892	294.8	1942	310.7		1992	356.42			
	1893	294.7	1943	310.5		1993	357.13			
	1894	294.8	1944	310.2		1994	358.61			
	1895	294.8	1945	310.3		1995	360.67			
	1896	294.9	1946	310.3		1996	362.58			
	1897	294.9	1947	310.4		1997	363.48			
	1898	294.9	1948	310.5		1998	366.27			
	1899	295.3	1949	310.9		1999	368.38			