数值计算复习笔记

徐大鹏

2017 年 6 月 14 日

1 概论

提到的考点：1.1 – 1.3.9 适定的、不适定的；误差来源和表示、条件数；机器精度；舍入误差：大数吃小数、抵消。

条件数：**解**的相对变化与**输入数据**相对变化的比值。

$$\text{condition number} = \frac{\left|\frac{(\hat{y}-y)}{y}\right|}{\left|\frac{(\hat{x}-x)}{x}\right|} = \left|\frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}}\right| = \left|\frac{x \cdot \Delta y}{y \cdot \Delta x}\right|$$

条件数刻画了问题的**病态性**，病态性的另一种说法是**敏感性**。也就是如果输入数据只变化了一点点，那么输出数据会变化多少。这个概念和在数学建模的灵敏性分析中使用的概念相一致。如果敏感性越强，意味着输入数据变化相同的范围，输出数据会有更大的变化，在上述定义式中，也就是分子更大，这样的结果就是条件数更大。

使用导数的概念可以得到条件数在某一点x的近似表示。

$$\text{condition number} = \left|\frac{x}{y} \cdot \frac{\Delta y}{\Delta x}\right| \overset{x \to \infty}{=} \left|\frac{x}{y} \cdot y'(x)\right| \overset{y=f(x)}{=} \left|\frac{x \cdot f'(x)}{f(x)}\right|$$

这种表示形式的最大好处在于，只和自变量x相关，便于分析和计算。上面这种形式的条件数是最常用的计算形式，必须熟记。

对于反函数$x = f^{-1}(y) = g(y)$而言，它的条件数是

$$\text{condition number} \overset{y \to \infty}{=} \left|\frac{y \cdot g'(y)}{x}\right| = \left|\frac{y \cdot \frac{1}{f'(x)}}{x}\right| = \left|\frac{y}{x \cdot f'(x)}\right| = \left|\frac{f(x)}{x \cdot f'(x)}\right|$$

正好就是原函数的条件数的倒数。

条件数的概念会和第二章线性方程组、第五章非线性方程组有联系。当x或y为零时，无法计算条件数，只能使用绝对条件数代替。

$$\text{absolute condition number} = \left|\frac{\Delta y}{\Delta x}\right|$$

浮点系统可以使用四个参数完备地表示：基数β、精度p、L、U。它们的含义是什么？

$$x = \pm\left(\sum_{i=0}^{p}\frac{d_i}{\beta^i}\right)\beta^E$$

正规化浮点系统的下溢限

$$\text{UFL} = \beta^L$$

上溢限

$$\text{OFL} = \beta^{U+1}(1-\beta^{-p})$$

正规化浮点数的总个数

$$2(\beta-1)\beta^{p-1}(U-L+1)+1$$

截断舍入的机器精度

$$\epsilon_{mach} = \beta^{1-p}$$

最近舍入的机器精度

$$\epsilon_{mach} = \frac{1}{2}\beta^{1-p}$$

机器精度的含义是什么？

舍入误差分析的标准模型：

$$\text{fl}(x \text{ op } y) = (x \text{ op } y)(1+\delta)$$

其中op可以是加、减、乘、除中的任意一种，相对扰动$|\delta| \leq \epsilon_{mach}$。

2 线性方程组

提到的考点：2.3 — 2.4.8 范数、性质、条件数、误差限、影响因素、残差；高斯消去法、LU分解。2.5 特殊线性方程组的解法：对称正定方程组的求解、带状方程组的求解。2.6 迭代法（参考第11章）：雅克比方法、高斯-赛德尔方法、SOR方法。

2.1 回顾：赋范线性空间

向量的p-范数：

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

注意表达式里，每一个分量x_i都取了绝对值。

矩阵范数：定义$m \times n$的矩阵$\boldsymbol{A}$的（向量诱导）范数是

$$\|\boldsymbol{A}\| = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$$

使用向量1-范数、∞-范数的结果不太容易推导，直接记住形式：

$$\|\boldsymbol{A}\|_1 = \max_{j} \sum_{i=1}^{m} |a_{ij}|$$

$$\|\boldsymbol{A}\|_\infty = \max_{i} \sum_{j=1}^{n} |a_{ij}|$$

一个是按行求和，一个是按列求和。

2.2 矩阵的条件数、误差限、残差

矩阵条件数的性质：

1. 对任意$\boldsymbol{A}$，$\mathrm{cond}(\boldsymbol{A}) \geq 1$。
   不会证。

2. 对任意$\boldsymbol{A}$及非零标量γ，$\mathrm{cond}(\gamma\boldsymbol{A}) = \mathrm{cond}(\boldsymbol{A})$。
   因为$\boldsymbol{I} = (\gamma\boldsymbol{A}) \cdot (\gamma\boldsymbol{A})^{-1} = (\gamma\boldsymbol{A}) \cdot (\gamma^{-1}\boldsymbol{A}^{-1})$，

$$\begin{aligned}
\mathrm{cond}(\gamma\boldsymbol{A}) &= \|\gamma\boldsymbol{A}\| \cdot \left\|\gamma^{-1}\boldsymbol{A}^{-1}\right\| \\
&= |\gamma| \cdot \|\boldsymbol{A}\| \cdot \left|\gamma^{-1}\right| \cdot \left\|\boldsymbol{A}^{-1}\right\| \\
&= \|\boldsymbol{A}\| \cdot \left\|\boldsymbol{A}^{-1}\right\| = \mathrm{cond}(\boldsymbol{A})
\end{aligned}$$

3. 对任意对角阵，$\boldsymbol{D} = \mathrm{diag}(d_i)$，$\mathrm{cond}(\boldsymbol{D}) = \frac{\max_i |d_i|}{\min_i |d_i|}$。
   证：显然。

条件数$\mathrm{cond}(\boldsymbol{A})$刻画了矩阵接近奇异的程度。行列式$\det(\boldsymbol{A})$刻画了矩阵是否是奇异的。

第2.3.4节主要讲了求解线性方程组的误差限表示、误差限和什么相关，以及如何推导误差限的问题。推导得到了两个主要结论。第一个是，右端向量带有扰动的方程组$\boldsymbol{A}(\boldsymbol{x}+\Delta\boldsymbol{x})=\boldsymbol{b}+\Delta\boldsymbol{b}$的解的估计式是

$$\frac{\|\Delta x\|}{\|x\|}\le \mathrm{cond}(\boldsymbol{A})\cdot\frac{\|\Delta\boldsymbol{b}\|}{\boldsymbol{b}}$$

第二个是，矩阵$\boldsymbol{A}$的元素带有扰动的方程组$(\boldsymbol{A}+\boldsymbol{E})(\boldsymbol{x}+\Delta\boldsymbol{x})=\boldsymbol{b}$的解的估计式是

$$\frac{\Delta\boldsymbol{x}}{\|\boldsymbol{x}+\Delta\boldsymbol{x}\|}\le \mathrm{cond}(\boldsymbol{A})\frac{\|\boldsymbol{E}\|}{\|\boldsymbol{A}\|}$$

注意直接使用不等式推得的第二个结论中，左边分母上是带扰动的输入$\boldsymbol{x}+\Delta\boldsymbol{x}$。

使用更精确的方法改进第二个结论？然后可以得到最终的结论：

$$\frac{\Delta\boldsymbol{x}}{\|\boldsymbol{x}\|}\le \mathrm{cond}(\boldsymbol{A})\left(\frac{\|\Delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}+\frac{\|\boldsymbol{E}\|}{\|\boldsymbol{A}\|}\right)$$

上式左侧的分式是向前误差，右侧的分式是向后误差。条件数是输出误差对输入误差的一个估计。一个输入数据的误差输入数据的误差就是数据传播误差。浮点系统的数据传播误差的大小由机器精度ϵ_{mach}决定。可以将上式写成这种形式

$$\frac{\Delta\boldsymbol{x}}{\|\boldsymbol{x}\|}\lessapprox \mathrm{cond}(\boldsymbol{A})\epsilon_{mach}$$

矩阵条件数$\mathrm{cond}(A)$太大的原因：

1. 矩阵本身是接近奇异的，也就是行向量之间接近线性相关；
2. 观测的数据相差太大，比如某个列向量大了很多个数量级

残差：

$$\boldsymbol{r}=\boldsymbol{b}-\boldsymbol{A}\hat{\boldsymbol{x}}$$

相对残差：

$$\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{A}\|\cdot\|\hat{\boldsymbol{x}}\|}$$

稳定算法产生的相对残差总是很小。

2.3 一般方法：Gauss消去和LU分解

线性方程组的求解：相互等价的两种算法，Gauss消去法和LU分解。 这里，我们构造了一个初等消去阵

$$\boldsymbol{M}_k=\boldsymbol{I}-\boldsymbol{m}\boldsymbol{e}_k^T$$

$$
\boldsymbol{M}_k\boldsymbol{a} = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & -m_{k+1} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -m_n & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ \boxed{a_k} \\ a_{k+1} \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}
$$

，其中

$$\boldsymbol{m} = (0, \cdots, 0, m_{k+1}, \cdots, m_n)$$

考虑到$m_i = a_i/a_k$，上式可写作

$$\boldsymbol{m} = (0, \cdots, 0, \frac{a_{i+1}}{a_k}, \cdots, \frac{a_n}{a_k})$$

$\boldsymbol{M}_k$是一个单位下三角矩阵，然后后面就很显然了，果然没什么好推的。

选主元：每次循环时选择最大的元素作为主元，因为主元a_k出现在$m_i = a_i/a_k$的分母上。如果a_k比较小，会产生一个非常大的乘数m_i，从而造成大数吃小数的（抵消）问题，导致误差。选主元的结果是在$\boldsymbol{A}$上乘以了一个排列矩阵$\boldsymbol{P}$，使

$$\boldsymbol{PA} = \boldsymbol{LU}$$

这样，求解$\boldsymbol{Ax} = \boldsymbol{b}$的问题等价转化为求解$\boldsymbol{PAx} = \boldsymbol{Pb}$的问题。特别注意，排列阵$\boldsymbol{P}$具有性质

$$P^{-1} = P^T$$

另外，这个性质实际上说明排列阵$\boldsymbol{P}$是一个正交矩阵。

高斯-若当（Gauss-Jordan）消去法：这种方法的区别在于，开始于

$$\left[\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{I} \end{array}\right]$$

但是行化简的方式，是将非对角线上的元素一次消去，是一种彻底的消元法。消元后，乘以一个对角矩阵将左侧矩阵的对角元素化为全1。最终得到的结果是

$$\left[\begin{array}{c|c} \boldsymbol{I} & \boldsymbol{A^{-1}} \end{array}\right]$$

这种方法更适合于求出逆矩阵$\boldsymbol{A}^{-1}$。

2.4 特殊方程组：Cholesky分解和简化的LU分解

对称正定方程组：如果$\boldsymbol{A}$是对称正定的，那么$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^T > 0$
楚列斯基分解的特点：

1. 所需计算的n个平方根都是正数
2. 不需要选主元就可以保证数值稳定
3. $\boldsymbol{A}$是对称正定的，上三角部分的元素用不到而且无需存储
4. 仅仅需要做$\frac{n^3}{6}$次乘法

Cholesky分解和LU分解的区别：

1. Cholesky分解得到的下三角矩阵$\boldsymbol{L}$不一定是单位下三角矩阵。
2. Cholesky分解不需要选主元
3. Cholesky分解需要计算平方根

最后，带状方程组的求解，是LU分解求解一般方程组的一个特殊情况。

3 线性最小二乘

提到的考点：正规方程组、增广方程组、QR分解（Householder变换、Givens旋转、Gram-Schmidt正交化方法）、奇异值分解。

3.1 目标、相关概念、惟一性

目标是最小化残差向量$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$。过程中最好选择2-范数，因为它与内积有关，具有正交性、光滑性、严格凸性。

一个一维函数f是**严格凸的**，如果这个函数在区间(x, y)上满足

$$\forall t \in (0, 1), \quad f(tx + (1-t)y) < tf(x) + (1-t)f(y)$$

范德蒙矩阵是列向量的每个分量上呈现等比关系的矩阵

$$V = \begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^{n-1} \\ 1 & \alpha_3 & \alpha_3^2 & \cdots & \alpha_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_m & \alpha_m^2 & \cdots & \alpha_m^{n-1} \end{bmatrix}$$

它的每一个元素是

$$\boldsymbol{V}_{i,j} = \alpha_i^{j-1}$$

当它是n阶方阵时，行列式为

$$\det(V) = \prod_{1 \le i < j \le n} (\alpha_j - \alpha_i)$$

线性最小二乘的解唯一的条件是，矩阵$\boldsymbol{A}$是列满秩的。

3.2 正规方程组

如果$\boldsymbol{A}$是列满秩的，则$n \times n$正定对称正规方程组$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}$与$m \times n$ 最小二乘问题$\boldsymbol{A}\boldsymbol{x} \cong \boldsymbol{b}$同解。使用Cholesky分解求解这个线性方程组的解。

正规方程组的敏感性很强，因为叉积矩阵$\boldsymbol{A}^T\boldsymbol{A}$的条件数是原矩阵条件数的平方。

3.3 QR分解

正交矩阵的特点：

1. 不会变化向量的2-范数。
2. $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{Q}\boldsymbol{Q}^{-1} = I$

课本先说了QR分解是什么，后说了如何做QR分解。QR分解的核心在于，对于每个$m \times n$矩阵$\boldsymbol{A}$能找到一个$m \times m$的正交矩阵$\boldsymbol{Q}$，使得

$$\boldsymbol{A} = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{bmatrix}$$

至于这个正交矩阵$\boldsymbol{Q}$是如何找到的，则是通过House-Hold变换、Givens旋转、Gram-Schmidt正交化等算法实现的。课本上所做的一个证明是，证明

$$\boldsymbol{Q}\begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{bmatrix}\boldsymbol{x} \cong \boldsymbol{b}$$

和

$$\begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{0} \end{bmatrix}\boldsymbol{x} \cong \boldsymbol{Q}^T\boldsymbol{b}$$

是同解的。这样，我们只需验证House-Hold变换等算法是实现了QR分解的

$$\boldsymbol{A} = \boldsymbol{Q}\begin{bmatrix}\boldsymbol{R}\\ \boldsymbol{0}\end{bmatrix}$$

就能说明算法能够正确求解出线性最小二乘了。

3.4 Householder变换

Householder变换：通过反射变换，把向量映射到坐标轴上，实现消元。课本上先讨论了特殊形式，我在这里整理的是House-Hold变换的一般形式。变换矩阵$\boldsymbol{H}$的形式是

$$\boldsymbol{H} = \boldsymbol{I} - 2\frac{\boldsymbol{v}\boldsymbol{v}^T}{\boldsymbol{v}^T\boldsymbol{v}}$$

在QR分解中，我们最终的目标是得到一个上三角矩阵$\boldsymbol{R}$。现在考虑每一个列向量是如何使用Householder变换消元的。对于一个m维列向量$\boldsymbol{a}$，考虑分块

$$\boldsymbol{a} = \begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{a}_2\end{bmatrix}$$

其中$\boldsymbol{a}_1$是$k-1$维向量$(1 \le k < m)$。我们的任务是什么呢？是要做QR分解，是要确定一个向量$\boldsymbol{v}$，它能将

$$\boldsymbol{H}(\boldsymbol{a}): \begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{a}_2\end{bmatrix} \mapsto \begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{0}\end{bmatrix} + \alpha\boldsymbol{e}_k$$

有了这个目标，我们看看如何确定那个$\boldsymbol{v}$。确定了$\boldsymbol{v}$也就确定了$\boldsymbol{H}$。一方面，

$$\begin{aligned}\boldsymbol{H}\boldsymbol{a} &= \left(\boldsymbol{I} - 2\frac{\boldsymbol{v}\boldsymbol{v}^T}{\boldsymbol{v}^T\boldsymbol{v}}\right)\boldsymbol{a}\\ &= \boldsymbol{a} - 2\boldsymbol{v}\frac{\boldsymbol{v}^T\boldsymbol{a}}{\boldsymbol{v}^T\boldsymbol{v}}\\ &= \begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{a}_2\end{bmatrix} - 2\boldsymbol{v}\frac{\boldsymbol{v}^T\boldsymbol{a}}{\boldsymbol{v}^T\boldsymbol{v}}\end{aligned}$$

另一方面，

$$\boldsymbol{H}(\boldsymbol{a}) = \boldsymbol{H}\boldsymbol{a} = \begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{0}\end{bmatrix} + \alpha\boldsymbol{e}_k$$

于是就有，

$$\begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{a}_2\end{bmatrix} - 2\boldsymbol{v}\frac{\boldsymbol{v}^T\boldsymbol{a}}{\boldsymbol{v}^T\boldsymbol{v}} = \begin{bmatrix}\boldsymbol{a}_1\\ \boldsymbol{0}\end{bmatrix} + \alpha\boldsymbol{e}_k$$

$$\begin{bmatrix} \mathbf{0} \\ \boldsymbol{a}_2 \end{bmatrix} - \alpha \boldsymbol{e}_k = 2\boldsymbol{v}\frac{\boldsymbol{v}^T\boldsymbol{a}}{\boldsymbol{v}^T\boldsymbol{v}}$$

$$\boldsymbol{v} = \frac{\boldsymbol{v}^T\boldsymbol{v}}{2\boldsymbol{v}^T\boldsymbol{a}}\left(\begin{bmatrix} \mathbf{0} \\ \boldsymbol{a}_2 \end{bmatrix} - \alpha \boldsymbol{e}_k\right)$$

为了计算简便，消去上式右端的标量系数是可行的。因为

$$\begin{aligned} H &= I - 2\frac{(\beta\boldsymbol{v})(\beta\boldsymbol{v})^T}{(\beta\boldsymbol{v})^T(\beta\boldsymbol{v})} \\ &= I - 2\frac{\beta^2\boldsymbol{v}\boldsymbol{v}^T}{\beta^2\boldsymbol{v}^T\boldsymbol{v}} \\ &= I - 2\frac{\boldsymbol{v}\boldsymbol{v}^T}{\boldsymbol{v}^T\boldsymbol{v}} \end{aligned}$$

所以$\boldsymbol{v}$的标量系数在$\boldsymbol{H}$的计算过程中会被消去，没有影响。这就得到了要取的$\boldsymbol{v}$：

$$\boldsymbol{v} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{a}_2 \end{bmatrix} - \alpha \boldsymbol{e}_k$$

其中

$$\alpha = -\mathrm{sgn}(a_k)\|\boldsymbol{a}_2\|_2$$

但是，这里的α为什么这么取值呢？$|\alpha| = \|a_2\|_2$的依据是什么呢？对于任意向量$\boldsymbol{a}$，正交矩阵$\boldsymbol{H}$作用于$\boldsymbol{a}$得到的线性映射$\boldsymbol{H}(\boldsymbol{a})$有着性质

$$\|\boldsymbol{a}\|_2 = \|\boldsymbol{H}\boldsymbol{a}\|_2$$

对于一个具体的向量

$$\boldsymbol{a} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{bmatrix}$$

根据上面的讨论，如果将线性映射的结果表示成

$$\boldsymbol{H}(\boldsymbol{a}) = \begin{bmatrix} \boldsymbol{a}_1 \\ \mathbf{0} \end{bmatrix} + \alpha \boldsymbol{e}_k$$

那么它一定满足性质

$$\left\| \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{bmatrix} \right\| = \left\| \begin{bmatrix} \boldsymbol{a}_1 \\ \mathbf{0} \end{bmatrix} + \alpha \boldsymbol{e}_k \right\|$$

根据勾股定理，

$$\left\|\begin{bmatrix}\boldsymbol{a}_1\\\boldsymbol{a}_2\end{bmatrix}\right\|_2^2=\|\boldsymbol{a}_1\|_2^2+\|\boldsymbol{a}_2\|_2^2$$

$$\left\|\begin{bmatrix}\boldsymbol{a}_1\\\boldsymbol{0}\end{bmatrix}+\alpha\boldsymbol{e}_k\right\|_2^2=\|\boldsymbol{a}_1\|_2^2+\alpha^2$$

因此

$$\alpha^2=\|\boldsymbol{a}_2\|_2^2$$

$$|\alpha|=\|\boldsymbol{a}_2\|$$

4 非线性方程组

提到的考点：5.1, 5.4 — 5.5.5：收敛速度、收敛性；二分法、牛顿法、反二次方法。迭代收敛准则。例题5.9

4.1 收敛速度和判停准则

第k步迭代的绝对误差是$\boldsymbol{e}_k=\boldsymbol{x}_k-\boldsymbol{x}^*$。考虑式子

$$\lim_{k\to\infty}\frac{\|\boldsymbol{e}_{k+1}\|}{\|\boldsymbol{e}_k\|^r}=C$$

其中，C是一个大于零的有限常数，r为迭代法的收敛速度。

两种判停准则：

1. 限定残差：
$$\left\|f(\boldsymbol{x}_k)\right\|<\epsilon$$
2. 限定相对误差：
$$\frac{\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_k\|}{\max\left\{\|\boldsymbol{x}_k\|,\boldsymbol{1}\right\}}<\epsilon$$

4.2 二分法

二分法使用的中点计算方式是

$$m=a+\frac{b-a}{2}$$

判断收缩区间的条件是

$$\operatorname{sgn}(f(a))=\operatorname{sgn}(f(m))$$

他们的优点是什么?（课本上有）

4.3 不动点迭代

迭代格式是

$$\boldsymbol{x}_{k+1} = g(\boldsymbol{x}_k)$$

收敛速度是线性的。迭代收敛的证明和收敛速度的证明，使用拉格朗日微分中值定理

$$\boldsymbol{e}_{k+1} = \boldsymbol{x}_{k+1} - \boldsymbol{x}^* = g(\boldsymbol{x}_k) - g(\boldsymbol{x}^*) = g'(\theta_k)(\boldsymbol{x}_k - \boldsymbol{x}^*) = g'(\theta_k)\boldsymbol{e}_k$$

如果需要证明平方收敛或者更高阶收敛，则使用泰勒中值定理。

4.4 牛顿法

迭代格式是

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \frac{f(\boldsymbol{x}_k)}{f'(\boldsymbol{x}_k)}$$

收敛速度在单根情形的收敛速度是二次的，在重根的收敛速度是线性的。

割线法略。

4.5 反二次插值

想法：改进割线法的一次函数，使用二次函数和x轴的交点横坐标作为下一个x_{k+1}。但是，二次函数可以与x轴没有交点，使用“反二次”就能保证一定有交点了。

为方便起见，分别记自变量x_{k-2}、x_{k-1}、x_k为a、b、c，他们对应的函数值分别为f_a、f_b、f_c。

构造反二次函数，即以y为自变量，x为因变量，

$$x = m_0 + m_1 y + m_2 y^2$$

在上式中令$y = 0$，可以求出曲线和x轴的交点

$$x_{k+1} = m_0$$

为确定m_0、m_1、m_2，使用拉格朗日插值公式：

$$x(y) = a\frac{(y-f_b)(y-f_c)}{(f_a-f_b)(f_a-f_c)} + b\frac{(y-f_a)(y-f_c)}{(f_b-f_a)(f_b-f_c)} + c\frac{(y-f_a)(y-f_b)}{(f_c-f_a)(f_c-f_b)}$$

取$y = 0$得到

$$x(0) = a\frac{f_b f_c}{(f_a-f_b)(f_a-f_c)} + b\frac{f_a f_c}{(f_b-f_a)(f_b-f_c)} + c\frac{f_a f_b}{(f_c-f_a)(f_c-f_b)}$$

5 插值

提到的考点：7.3 三种基函数：单项式、拉格朗日、牛顿。7.3.1 — 7.3.3, 7.3.5 余项定理。7.4.2 三次样条。例题7.6

5.1 牛顿多项式插值

对给定的数据点集(t_j, y_j)牛顿基函数

$$\pi_j(t) = \prod_{k=1}^{j-1}(t - t_k), \quad j = 1, \cdots, n$$

对应的插值多项式

$$p_{n-1}(t) = \sum_{j=1}^{n} x_j \pi_j(t) = \sum_{j=1}^{n} \left[x_j \prod_{k=1}^{j-1}(t - t_k) \right]$$

为了求这个多项式，对应的基底矩阵$\boldsymbol{A}$是

$$\begin{bmatrix} 1 & & \cdots & & 0 \\ 1 & t_1 - t_0 & & & \\ 1 & t_2 - t_0 & (t_2 - t_0)(t_2 - t_1) & & \vdots \\ \vdots & \vdots & & \ddots & \\ 1 & t_k - t_0 & \cdots & \cdots & \prod_{j=0}^{n-1}(t_n - t_j) \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ \\ x_n \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ \\ y_n \end{bmatrix}$$

这是一个下三角矩阵，每一个x_j由当前行和之前的所有行确定。这说明，当新增加一个数据点t_{n+1}, y_{n+1}时，可以直接确定系数x_{n+1}的值，并且在p_{n-1}中增加一项

$$x_{n+1} = \frac{y_{n+1} - p_n(t_{n+1})}{\pi_{n+1}(t_{n+1})} = \frac{y_{n+1} - p_n(t_{n+1})}{\prod_{k=1}^{n}(t_{n+1} - t_k)}$$

就可以得到p_n。

6 积分和微分

提到的考点：待定系数法。8.3 牛顿科特斯方法、高斯方法。简单求积公式构造出复杂的求积公式。8.6 差分公式和推导。积分：代数精度、验证法确定。