

化工应用数学 第四章 方程(组)数值求解 讲义

02.非线性方程

非线性的现实:

上一节我们主要讲述了线性方程组的求解以及数值解法,虽然在计算机图像处理等问题上线性方程组有很多用途,但在实际工业应用中,我们所遇到的往往是非线性的问题。

比如气体的状态方程,我们在中学所学习的往往是 $PV=nRT$ 这样一个看起来是线性的问题,但是在更高层次的热力学等学习阶段,我们就会发现这一方程实际上是一种简化与近似,而更精确的则是下面这种非线性的形式:

$$\left(P + \frac{an^2}{V^2}\right)(V - nb) = nRT$$

对不同气体的饱和蒸气压与温度的关联式往往也是呈现为非线性的形式;同样的,非线性现象的两个典型的例子还有蝴蝶效应以及双摆的轨迹。

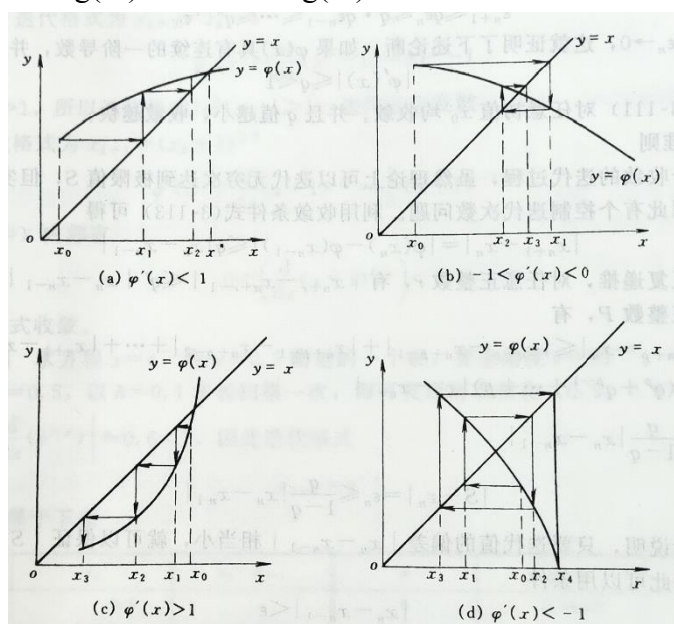
蝴蝶效应:线性关系的话,输出要增大多少倍,输入也必须增大相应的倍数。

双摆:轴互相平行,一个摆的支点装在另一摆的下部所形成的组合物体。

非线性方程求解-迭代:

对非线性方程,仅在次数不大于 4 的时候有求解公式,同时对于一元四次方程解也会十分的复杂,所以对于较高次数的非线性方程我们一般是没有相应的求解公式的,所以我们一般会使用迭代的方法对其进行求解。

迭代求解与之前线性方程组迭代求解的思想相同,都是根据已知方程与相关的数学物理知识构建适合的迭代公式,如 $x=g(x)$,选择合适的初值 x^0 ,之后逐步迭代: $x^1=g(x^0)$ 、 $x^2=g(x^1)$ 、...、 $x^{k+1}=g(x^k)$ 、... 直至达到精度要求



迭代的直观解释与收敛性:

若迭代公式是由已知函数 $f(x)$ 变换过来的, 那么迭代公式 $x=g(x)$ 就有了明确的意义, 即方程组的解为 $y=x$ 和 $y=g(x)$ 的交点处的 x 值。(如上图所示)

那么, 迭代所代表的过程如下: 首先, 从初值 x_0 做 x 轴垂线与 $g(x)$ 相交于 $(x_0, g(x_0))$ (此步骤对应于 $y=g(x)$), 之后通过 $(x_0, g(x_0))$ 做 x 轴平行线与 $y=x$ 相交于 $(g(x_0), g(x_0))$ (此步骤对应于 $y=x$), 再然后通过 $(x_1=g(x_0), g(x_0))$ 做 x 轴垂线与 $g(x)$ 相交于 $(x_1, g(x_1))$ (此步骤对应于 $y=g(x)$), ...

迭代的过程就是不断重复上面步骤的过程。

从上图中可以看出, 针对不同的情况, 当函数 $f(x)$ 的导数绝对值小于 1 时, 迭代收敛, 否则不收敛。

思考: 若函数 $f(x)$ 的导数绝对值在定义域内既有大于 1 也有小于 1 的情况, 会有什么结论?

二分法:

二分法可以看作是一个特殊的迭代方法, 因为在构造迭代公式的时候, 没有使用原函数 $f(x)$, 而是利用数学上的介值定理:

**$f(x)$ 在 $[a, b]$ 上连续, M 是其最大值, m 是其最小值,
则对任意的 $y, m \leq y \leq M$, 至少存在一点 $\xi \in [a, b]$, 使得
 $f(\xi)=y$**

或者说是介值定理的一个特殊情况, 零点存在定理:

若函数 $y = f(x) \in C([a, b])$, 且 $f(a) \cdot f(b) < 0$, 则至少存在一点 $x_0 \in (a, b)$, 使得 $f(x_0) = 0$ 。

其一般的步骤如下:

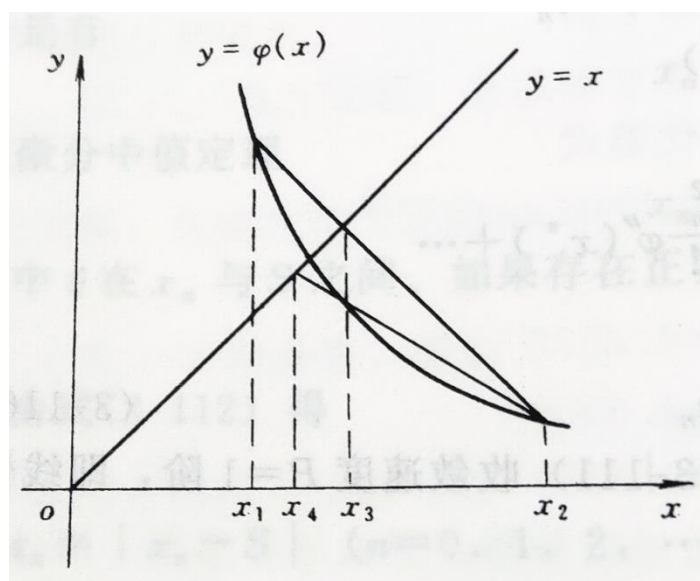
1. 求该区间的中点 $m=(a+b)/2$, 并找出 $f(m)$ 的值
2. 若 $f(m)$ 与 $f(a)$ 正负号相同, 则取 $[m, b]$ 为新的区间, 否则取 $[a, m]$
3. 重复上述两步, 直到得到理想的精确度为止。

思考: 1. 二分法的迭代公式是什么? 2. 二分法的精度和收敛性如何? 3. 若存在多个解, 如何全部求出来?

威格斯坦(Wegstein)法:

一般的迭代法受到导数绝对值必须小于 1 的限制, 而且收敛速度较慢, 针对这些问题, 学者们提出了很多不同的改进迭代方法, 威格斯坦法就是其中之一。

与一般迭代法中, 从一初值出发, 在函数 $y=g(x)$ 和 $y=x$ 之间不断做折线不同, 威格斯坦法采用两个初始值 x_1 、 x_2 , 利用 $(x_1, g(x_1))$ 和 $(x_2, g(x_2))$ 连线与 $y=x$ 交点的 x 值作为新值 x_3 , 之后再利用 x_2 、 x_3 作为初值重复上面的步骤, 具体步骤见下图。



通过数学推导，我们可以得到威格斯坦法的迭代公式为：

$$x_{n+1} = \frac{x_n \varphi(x_{n-1}) - x_{n-1} \varphi(x_n)}{(x_n - x_{n-1}) - [\varphi(x_n) - \varphi(x_{n-1})]}$$

若记斜率为 S ，同时有 $C=1/(1-S)$ ，那么上式可以变为：

$$x_{n+1} = (1-C)x_n + C\varphi(x_n)$$

令 $q=1-C$ ，可以看出，当 $C=1$ 即 $q=0$ 时，上式与一般迭代法是相同的，即 $C=1$ 时威格斯坦法退化为简单迭代法。

研究发现， $0 < q < 1$ 时，迭代为有阻尼的顺序迭代法，通常能够稳定收敛，但收敛速度较慢； $q < 0$ 时收敛速度较快，但容易导致不稳定；一般计算过程中可以取 $-5 < q < 0$ ，此时称为有界的威格斯坦法。

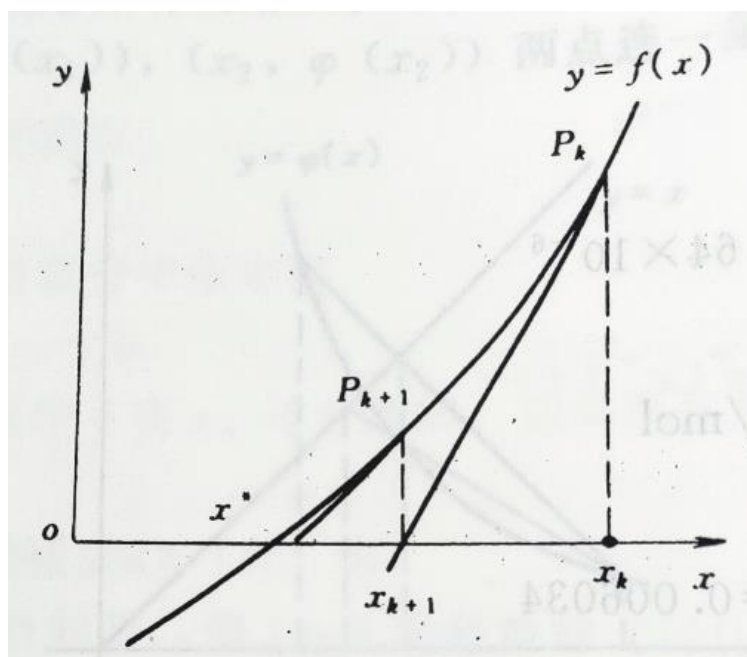
牛顿法：

通过分析可以发现威格斯坦法本身还是有一定的限制，对于一些能够解析求导的方程，我们还有一种特别有效的方法，即牛顿法，其特点是程序实现简单，并且只要选择适当的初值，收敛速度也很快。

牛顿法的基本思想是把非线性方程线性化，用线性方程的解去逐步逼近非线性方程的解。

牛顿法的基本步骤如下：

1. 选取某一初值 x_0 （初始预测值）；
2. 通过 $(x_0, f(x_0))$ 作函数 $y=f(x)$ 切线；
3. 以切线与 x 轴交点 x_1 作为新的预测值；
4. 重复上述步骤直至达到精度要求。



针对上图，根据导数/斜率的定义可知：

$$f'(x_k) = \frac{f(x_k)}{x_k - x_{k+1}}$$

即：

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

上式即为牛顿法的迭代公式。

牛顿法的优点是收敛速度快，在根附近具有二阶收敛特性；但对某些情况，如果初值选择不合适的话，可能会导致收敛困难甚至是发散。

同时，在后续深度神经网络的课程中我们也会发现基于牛顿法的一个变种方式：梯度下降法，在那里引入了学习率， α ：

$$x_{k+1} = x_k - \alpha \frac{f(x_k)}{f'(x_k)}$$