2016 03 29

Chapter 7:

§7.1

We are solving the ODE

$$\vec{y}'(t) = \vec{f}(t, \vec{y}(t)), \quad t \geq 0, \quad \vec{y}(t_0) = \vec{y}_0.$$

Assume $\vec{f}$ is non-linear.

The backward Euler method is

$$\vec{y}_{n+1} = \vec{y}_n + h\vec{f}(t_{n+1}, \vec{y}_{n+1}).$$

The trapezoidal rule is

$$\vec{y}_{n+1} = \vec{y}_n + \tfrac{1}{2}h(\vec{f}(t_{n+1}, \vec{y}_{n+1}) + \vec{f}(t_n, \vec{y}_n)).$$

The general form of these two methods can be written as

$$\vec{\omega} = h\vec{g}(\vec{\omega}) + \vec{\beta}, \quad \vec{\omega} \in \mathbb{R}^d. \qquad \text{——— (*)}$$

For the backward Euler method, we have

$$\vec{g}(\vec{\omega}) = \vec{f}(t_{n+1}, \vec{\omega}), \quad \vec{\beta} = \vec{y}_n.$$

For the trapezoidal rule, we have

$$\vec{g}(\vec{\omega}) = \tfrac{1}{2}\vec{f}(t_{n+1}, \vec{\omega}), \quad \vec{\beta} = \vec{y}_n + \tfrac{1}{2}h\vec{f}(t_n, \vec{y}_n).$$

Implicit Runge-Kutta methods and multi-step methods may also be written in this form.

Goal: Given some initial guess $\vec{\omega}^{[0]}$, find algorithm

$$\vec{\omega}^{[i+1]} = \vec{s}(\vec{\omega}^{[i]}), \quad i = 0, 1, 2, \ldots$$

such that:

1) $\vec{\omega}^{[i]} \to \vec{\omega}$ which is solution to (*);
2) cost is small;
3) algorithm converges fast.

The most simple is fixed-point iteration (also known as functional iteration sometimes):

$$\vec{\omega}^{[i+1]} = h\vec{g}(\vec{\omega}^{[i]}) + \vec{\beta}, \quad i = 0, 1, 2, \ldots$$

Notation: $\|\cdot\|$ is a vector norm and $B_\rho(\vec{\omega})$ is a closed ball of radius $\rho > 0$ centered at $\vec{\omega}$:

$$B_\rho(\vec{\omega}) = \{\vec{u} \in \mathbb{R}^d ; \|\vec{u} - \vec{\omega}\| \leq \rho\}.$$

Theorem 7.1: Let $h>0$, $\vec{w}^{[0]} \in \mathbb{R}^d$, and suppose there are $\lambda \in (0,1)$ and $\rho > 0$ such that:

(i) $\|\vec{g}(\vec{v}) - \vec{g}(\vec{u})\| \le \frac{\lambda}{h}\|\vec{v}-\vec{u}\| \quad \forall \vec{v},\vec{u} \in \mathcal{B}_\rho(\vec{w}^{[0]})$;

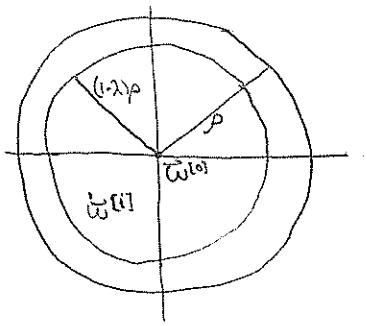
(ii) $\vec{w}^{[1]} \in \mathcal{B}_{(1-\lambda)\rho}(\vec{w}^{[0]})$.

Then:

(a) $\vec{w}^{[i]} \in \mathcal{B}_\rho(\vec{w}^{[0]})$ (solution does not diverge);

(b) $\hat{w} = \lim_{i\to\infty} \vec{w}^{[i]}$ exists and obeys $\hat{w} = h\vec{g}(\hat{w}) + \vec{\beta}$;

(c) $\hat{w}$ is unique.



Proof: We prove first that

$$\|\vec{w}^{[i+1]} - \vec{w}^{[i]}\| \le \lambda^i(1-\lambda)\rho.$$

This holds for $i=0$ by (ii). We proceed by induction. Assume the statement is true $\forall m = 0,1,\dots,i-1$. Then

$$\begin{aligned}\|\vec{w}^{[i+1]} - \vec{w}^{[i]}\| &= \left\|\left(h\vec{g}(\vec{w}^{[i]}) + \vec{\beta}\right) - \left(h\vec{g}(\vec{w}^{[i-1]}) + \vec{\beta}\right)\right\| \\ &= h\|\vec{g}(\vec{w}^{[i]}) - \vec{g}(\vec{w}^{[i-1]})\| \\ &\le \lambda\|\vec{w}^{[i]} - \vec{w}^{[i-1]}\| \qquad \text{(by (i))} \\ &\le \lambda\lambda^{i-1}(1-\lambda)\rho \\ &= \lambda^i(1-\lambda)\rho.\end{aligned}$$

↑ don't know (i) is applicable

This completes the induction.

this should all be one single induction

Note that

$$\vec{w}^{[i+1]} - \vec{w}^{[0]} = \sum_{j=0}^{i}\left(\vec{w}^{[j+1]} - \vec{w}^{[j]}\right), \quad i = 0,1,2,\dots$$

and so

$$\begin{aligned}\|\vec{w}^{[i+1]} - \vec{w}^{[0]}\| &= \left\|\sum_{j=0}^{i}\left(\vec{w}^{[j+1]} - \vec{w}^{[j]}\right)\right\| \\ &\le \sum_{j=0}^{i}\|\vec{w}^{[j+1]} - \vec{w}^{[j]}\|\end{aligned}$$

$$\leq \sum_{j=0}^{i} \lambda^j (1-\lambda)\rho$$
$$= (1-\lambda^{i+1})\rho$$
$$\leq \rho.$$

This proves (a).

Now we turn to (b). Note

$$\vec{\omega}^{[i+k]} - \vec{\omega}^{[i]} = \sum_{j=0}^{k-1} (\vec{\omega}^{[i+j+1]} - \vec{\omega}^{[i+j]})$$

and so

$$\|\vec{\omega}^{[i+k]} - \vec{\omega}^{[i]}\| \leq \sum_{j=0}^{k-1} \|\vec{\omega}^{[i+j+1]} - \vec{\omega}^{[i+j]}\|$$
$$\leq \sum_{j=0}^{k-1} \lambda^{i+j}(1-\lambda)\rho$$
$$= \lambda^i (1-\lambda^k)\rho.$$

bad wording

Since $\lambda \in (0,1)$, there exists k large enough such that for any $\varepsilon > 0$

$$\|\vec{\omega}^{[i+k]} - \vec{\omega}^{[i]}\| < \varepsilon.$$

Therefore $(\vec{\omega}^{[i]})_{i=0}^{\infty}$ is a Cauchy sequence. Since $B_\rho(\vec{\omega}^{[0]})$ is complete, there is $\hat{\omega} \in B_\rho(\vec{\omega}^{[0]})$ such that $\|\hat{\omega} - \vec{\omega}^{[i]}\| \to 0$ as $i \to \infty$. This proves (b).

We prove (c) by contradiction. Assume $\vec{\omega}^* \in B_\rho(\vec{\omega}^{[0]})$ with $\vec{\omega}^* \neq \hat{\omega}$ and $\vec{\omega}^* = h\vec{g}(\vec{\omega}^*) + \vec{\beta}$. Then $\|\vec{\omega}^* - \hat{\omega}\| > 0$ implies

$$\|\vec{\omega}^* - \hat{\omega}\| = \|(h\vec{g}(\vec{\omega}^*) + \vec{\beta}) - (h\vec{g}(\hat{\omega}) + \vec{\beta})\|$$
$$= h\|\vec{g}(\vec{\omega}^*) - \vec{g}(\hat{\omega})\|$$
$$\leq \lambda \|\vec{\omega}^* - \hat{\omega}\|$$
$$< \|\vec{\omega}^* - \hat{\omega}\|.$$

■

2016 03 31

Let's look at condition (i) of theorem 7.1. Let's assume that $\vec{g}$ is smoothly differentiable. Then there is $\tau \in (0,1)$ such that

$$\vec{g}(\vec{v}) - \vec{g}(\vec{u}) = \frac{\partial \vec{g}(\tau\vec{v} + (1-\tau)\vec{u})}{\partial \vec{\omega}} (\vec{v} - \vec{u}).$$

Hence

$$\|\vec{g}(\vec{v}) - \vec{g}(\vec{u})\| \leq \left\| \frac{\partial \vec{g}(\tau\vec{v} + (1-\tau)\vec{u})}{\partial \vec{\omega}} \right\| \|\vec{v} - \vec{u}\|.$$

So (i) says something about the magnitude of h in relation to

$$\left\| \frac{\partial \vec{g}}{\partial \vec{\omega}} \right\|.$$

For the backward Euler method, we have

$$\left\| \frac{\partial \vec{g}}{\partial \vec{\omega}} \right\| = \left\| \frac{\partial \vec{f}}{\partial \vec{\omega}}(t_{n+1}, \vec{y}_{n+1}) \right\| \leq \frac{\lambda}{h},$$

and so (i) gives

$$h \left\| \frac{\partial \vec{f}}{\partial \vec{\omega}} \right\| \leq \lambda < 1$$

which means

$$h < \left\| \frac{\partial \vec{f}}{\partial \vec{\omega}} \right\|^{-1}.$$

So the fixed point iteration imposes a restriction on h. Can we do better?

§7.2 Newton-Raphson method

We are trying to solve $\vec{\omega} = h\vec{g}(\vec{\omega}) + \beta$. Note

$$\vec{\omega} = h\vec{g}(\vec{\omega}^{[i]} + \vec{\omega} - \vec{\omega}^{[i]}) + \vec{\beta}$$

$$= \vec{\beta} + h\vec{g}(\vec{\omega}^{[i]}) + h\frac{\partial \vec{g}}{\partial \vec{\omega}}(\vec{\omega}^{[i]})(\vec{\omega} - \vec{\omega}^{[i]}) + \mathcal{O}(\|\vec{\omega} - \vec{\omega}^{[i]}\|^2).$$

Therefore

$$\vec{\omega} - \vec{\omega}^{[i]} \approx \vec{\beta} + h\vec{g}(\vec{\omega}^{[i]}) - \vec{\omega}^{[i]} + h\frac{\partial \vec{g}}{\partial \vec{\omega}}(\vec{\omega}^{[i]})(\vec{\omega} - \vec{\omega}^{[i]}).$$

This suggests the method

$$\underbrace{\left(I - h\frac{\partial \vec{g}}{\partial \vec{\omega}}(\vec{\omega}^{[i]})\right)}_{A}\underbrace{(\vec{\omega} - \vec{\omega}^{[i]})}_{u} \approx \underbrace{\vec{\beta} + h\vec{g}(\vec{\omega}^{[i]}) - \vec{\omega}^{[i]}}_{F}.$$

The Newton-Raphson method is

$$\vec{\omega}^{[i+1]} = \vec{\omega}^{[i]} - \left(I - h\frac{\partial \vec{g}}{\partial \vec{\omega}}(\vec{\omega}^{[i]})\right)^{-1}\left(\vec{\omega}^{[i]} - \vec{\beta} - h\vec{g}(\vec{\omega}^{[i]})\right).$$

Properties of Newton-Raphson:

1) For $h > 0$ small enough, Newton-Raphson has quadratic convergence:

$$\|\vec{\omega}^{[i+1]} - \hat{\omega}\| \leq c\|\vec{\omega}^{[i]} - \hat{\omega}\|^2.$$

2) At every iteration, we need to compute $\frac{\partial \vec{g}}{\partial \vec{w}}$, which is very expensive.
3) At every iteration, we must solve $AU = F$, which is also very expensive.

Modified Newton-Raphson:
Define

$$J = \frac{\partial \vec{g}}{\partial \vec{w}}(\vec{w}^{(0)})$$

Our new scheme is

$$\vec{w}^{[i+1]} = \vec{w}^{[i]} - (I - hJ)^{-1}(\vec{w}^{[i]} - \vec{\beta} - h\vec{g}(\vec{w}^{[i]})), \quad i = 0, 1, 2, \ldots .$$

This is nice because we need to compute J only once and because we can reuse $(I-hJ)^{-1}$. However, we lose quadratic convergence.
Note that the modified Newton-Raphson method is a fixed point method:

$$\vec{w}^{[i+1]} = \vec{w}^{[i]} - (I-hJ)^{-1}(\vec{w}^{[i]} - \vec{\beta} - h\vec{g}(\vec{w}^{[i]}))$$
$$\Rightarrow (I-hJ)(\vec{w}^{[i+1]} - \vec{w}^{[i]}) = -\vec{w}^{[i]} + \vec{\beta} + h\vec{g}(\vec{w}^{[i]})$$
$$\Rightarrow (I-hJ)\vec{w}^{[i+1]} - \vec{w}^{[i]} + hJ\vec{w}^{[i]} = -\vec{w}^{[i]} + \vec{\beta} + h\vec{g}(\vec{w}^{[i]})$$
$$\Rightarrow (I-hJ)\vec{w}^{[i+1]} = h(\vec{g}(\vec{w}^{[i]}) - J\vec{w}^{[i]}) + \vec{\beta}$$
$$\Rightarrow \vec{w}^{[i+1]} = h(I-hJ)^{-1}(\vec{g}(\vec{w}^{[i]}) - J\vec{w}^{[i]}) + \vec{\beta}.$$

This is a fixed point method where

$$\tilde{g}(\vec{w}) = (I-hJ)^{-1}(\vec{g}(\vec{w}) - J\vec{w}),$$
$$\tilde{\beta} = (I-hJ)^{-1}\vec{\beta}.$$

Applying theorem 7.1, we have convergence.
Recall the assumptions of theorem 7.1. If g is Lipschitz then (i) is satisfied:

$$\|\vec{g}(\vec{v}) - \vec{g}(\vec{u})\| \leq L\|\vec{v} - \vec{u}\|$$
$$\frac{1}{h}\|\vec{u} - \vec{v}\| = L\|\vec{u} - \vec{v}\|$$

As $h \to 0$ we need $\|\vec{u} - \vec{v}\| \to 0$, $\rho \to 0$.