Practical enhancements for Newton-type methods

'Soft' tangent may lead to divergence.

$$\tilde{K}\,\Delta d^{(i)} = F^{ext} - K d^{(i)} = R^{(i)}$$

approximation of K

search direction

We are thinking of a linear system here.

$$d^{(i+1)} = d^{(i)} + s^{(i)}\,\Delta d^{(i)}$$

line search parameter.

$s^{(i)}$ gives the incremental size in the direction of $\Delta d^{(i)}$.

Two approaches to determine $s^{(i)}$

1.
$$P(d) := \frac{1}{2} d^T K d - d^T F^{ext}$$

$$= \frac{1}{2} K_{PQ}\, d_P d_Q - d_P F_P^{ext}$$

$$\frac{\partial P}{\partial d_R} = \frac{1}{2} K_{PQ}\,\delta_{PR}\, d_Q + \frac{1}{2} K_{PQ}\, d_P\, \delta_{QR} - \delta_{PR} F_P^{ext}$$

$$= \frac{1}{2} K_{RQ}\, d_Q + \frac{1}{2} K_{PR}\, d_P - F_R^{ext}$$

$$= K_{RP}\, d_P - F_R^{ext}$$

$\Rightarrow$ P is minimized at $\frac{\partial P}{\partial d} = 0$, which is $Kd = F^{ext}$.

$$\varphi(s^{(i)}) = P(d^{(i)} + s^{(i)} \Delta d^{(i)})$$

We choose $s^{(i)}$ s.t. φ is minimized: $\dfrac{d\varphi}{ds} = 0$

$$0 = \frac{d}{ds}\left\{ \frac{1}{2}(d^{(i)} + s\,\Delta d^{(i)})^T K (d^{(i)} + s\,\Delta d^{(i)}) - (d^{(i)} + s\,\Delta d^{(i)})^T F^{ext} \right\}$$

$$= \frac{1}{2}\Delta d^{(i)T} K (d^{(i)} + s\,\Delta d^{(i)}) + \frac{1}{2}(d^{(i)} + s\,\Delta d^{(i)}) K\, \Delta d^{(i)}$$

$$- \Delta d^{(i)T} F^{ext}$$

$$\Rightarrow \quad 0 = \Delta d^{(i)T} K d^{(i)} + s\, \Delta d^{(i)T} K\, \Delta d^{(i)} - \Delta d^{(i)T} F^{ext}.$$

$$\Rightarrow \quad \left(\Delta d^{(i)T} K\, \Delta d^{(i)}\right) s = \Delta d^{(i)T} (F^{ext} - K d^{(i)})$$

$$= \Delta d^{(i)T} R^{(i)}$$

$$\Rightarrow \quad s^{(i)} = \frac{\Delta d^{(i)T} R^{(i)}}{\Delta d^{(i)T} K\, \Delta d^{(i)}}$$

$$\frac{d^2\varphi}{ds^2} = \Delta d^{(i)T} K\, \Delta d^{(i)} > 0 \quad \text{if } K \text{ is positive-definite.}$$

Verify!

meaning $s^{(i)}$ indeed minimizes the potential P.

2. idea: select $s^{(i)}$ such that $R^{(i+1)} := F^{ext} - K(d^{(i)} + s^{(i)} \Delta d^{(i)})$ has zero component in the direction of $\Delta d^{(i)}$:

$$\Delta d^{(i)^T} R^{(i+1)} = 0$$

This strategy is more general as we do not need a potential P.

Now we apply the idea to nonlinear problems:

$$\hat{K} \Delta d^{(i)} = F^{ext} - N(d^{(i)})$$

$$d^{(i+1)} = d^{(i)} + s^{(i)} \Delta d^{(i)}$$

to determine $s^{(i)}$, we define

$$G(s^{(i)}) := \Delta d^{(i)^T} R^{(i+1)}$$

$$= \Delta d^{(i)^T} \left(F^{ext} - N(d^{(i)} + s^{(i)} \Delta d^{(i)}) \right)$$

Our design: $G(s^{(i)}) = 0$

$\rightarrow$ a scalar nonlinear problem.

it can be computationally intensive

We thus release this condition to $\underline{G(s^{(i)}) \approx 0.}$

$$| G(s^{(i)}) | \leq \frac{1}{2} | G(0) |$$

Reference : H. Matthies & G. Strang , IJNME 14 : 1613-1626, 1979.

Remark 1 : For nonlinear elasticity, there is a potential

$$U(d) \text{ s.t. } N(d) = \frac{\partial U}{\partial d} .$$

$$P(d) = U(d) - d^T F^{ext}$$

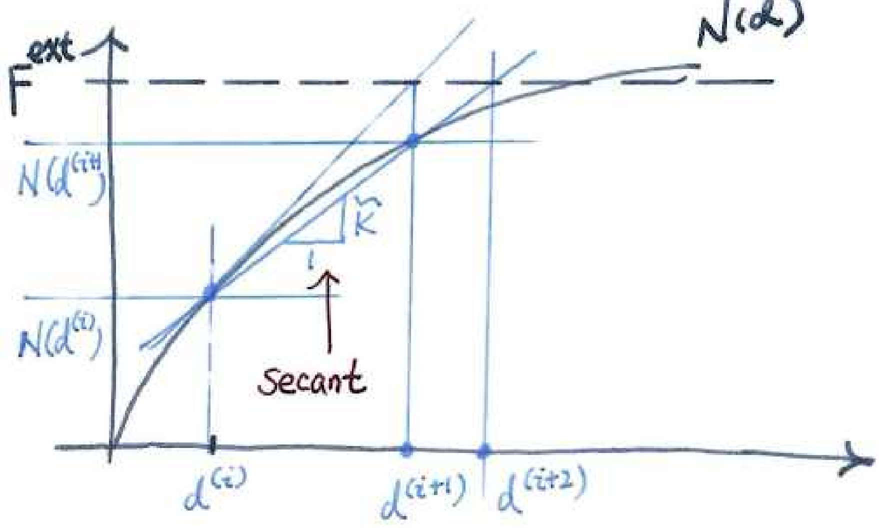
The first approach can be applied !

Remark 2 : minimizing $R^{(i+1)^T} R^{(i+1)}$ is an alternate option.

Remark 3 : We typically limit $|s^{(i)}|$ to be less than 1.

It is intended to be used as an insurance to prevent divergence due to the 'soft' mode.

Quasi-Newton methods



- $d^{(i)}$ & $d^{(i+1)}$ are obtained.

- $\Delta R^{(i)} = R^{(i+1)} - R^{(i)} = -\left(N(d^{(i+1)}) - N(d^{(i)})\right)$

- Secant: $\tilde{K} := \dfrac{-\Delta R^{(i)}}{d^{(i+1)} - d^{(i)}}$ $\quad (*)$

  and $\tilde{K}(d^{(i+2)} - d^{(i+1)}) = R^{(i+1)} = F^{ext} - N(d^{(i+1)})$

  → $\tilde{K}$ is used to determine $d^{(i+2)}$

We need to generalize the definition $(*)$ for multi-dof problems.

$$\tilde{K}(d^{(i+1)} - d^{(i)}) = -\Delta R^{(i)}$$

→ Quasi-Newton equation: a design criterion for multi-dof problems.

Broyden-Fletcher-Goldfarb-Shanno (BFGS)

$$\tilde{K}^{-1} := (I + v w^T)\,\bar{K}^{-1}(I + w v^T)$$

Design criteria:

a) Quasi-Newton equation;

b) $\bar{K}^{-1}$ symmetry implies $\tilde{K}^{-1}$ symmetry;

c) $\bar{K}^{-1} > 0$ & $v^T w \neq -1$ implies $\tilde{K}^{-1} > 0$.

$$\boxed{\begin{array}{l}\text{Quasi-Newton + Line searches:} \\ \qquad \tilde{K}\, s^{(i)} \Delta d^{(i)} = -\Delta R^{(i)}\end{array}}$$

or, equivalently $\underline{\tilde{K}^{-1} \Delta R^{(i)} = - s^{(i)} \Delta d^{(i)}}$

Choice 1: $v := \dfrac{\Delta d^{(i)}}{\Delta d^{(i)} \cdot \Delta R^{(i)}}$ $\quad w := -\Delta R^{(i)} + \alpha^{(i)} R^{(i)}$

$$\alpha^{(i)} := \left(\frac{-s^{(i)} \Delta R^{(i)} \cdot \Delta d^{(i)}}{R^{(i)} \cdot \Delta d^{(i)}} \right)^{1/2}$$

Claim: If $\bar{K} \Delta d^{(i)} = R^{(i)}$, then Choice 1 makes the Quasi-Newton eqn. satisfied.

proof: $$(I + w v^T)\, \Delta R^{(i)} = \Delta R^{(i)} + v \cdot \Delta R^{(i)}\, w$$
$$= \alpha^{(i)} R^{(i)}$$

$$\bar{K}^{-1} \alpha^{(i)} R^{(i)} = \alpha^{(i)} \Delta d^{(i)}$$

$$(I + v w^T)\, \alpha^{(i)} \Delta d^{(i)} = \alpha^{(i)} \Delta d^{(i)} + \left(- \Delta R^{(i)} \cdot \Delta d^{(i)} \alpha^{(i)} + (\alpha^{(i)})^2 R^{(i)} \cdot \Delta d^{(i)} \right) \frac{\Delta d^{(i)}}{\Delta d^{(i)} \cdot \Delta R^{(i)}}$$
$$= \left(- \frac{s^{(i)}\, \Delta R^{(i)} \cdot \Delta d^{(i)}}{R^{(i)} \cdot \Delta d^{(i)}}\, R^{(i)} \cdot \Delta d^{(i)} \right) \frac{\Delta d^{(i)}}{\Delta d^{(i)} \cdot \Delta R^{(i)}}$$

$$= - s^{(i)} \Delta d^{(i)}.$$

$\blacksquare$

$$v \cdot w = -1 + \underbrace{\alpha^{(i)} R^{(i)} \cdot \Delta d^{(i)} \Big/ \Delta d^{(i)} \cdot \Delta R^{(i)}}_{\text{can become small}} \Rightarrow \tilde{K} \text{ is nearly singular}$$

monitor the eigenvalue of $I + vw^T$, if dangerous update arise, skip the update and reuse the matrix.

a total of i rank-2 updates.

$$\tilde{K}^{-1} = (I + v^{(i)} w^{(i)T})(I + v^{(i-1)} w^{(i-1)T}) \cdots (I + v^{(1)} w^{(1)T})\, \bar{K}^{-1} (I + w^{(1)} v^{(1)T})(I + v^{(2)} w^{(2)T}) \cdots (I + w^{(i)} v^{(i)T}) \quad \bigstar$$

In ★, $\bar{K}$ is the most recent formed & factored matrix

Remark 1 : K^{-1} is the action of K^{-1} on a vector,

or it represents the capability of obtaining a solution from the equation of $Kd = F$.

Remark 2 : Solving $\bar{K}\, \Delta d^{(i+1)} = R^{(i+1)}$ in the iterate $i+1$ is achieved in 3 steps

Step one : right-side updates

$$\bar{R}^{(1)} = R^{(i+1)} + v^{(i)} \cdot R^{(i+1)}\, w^{(i)}$$

$$\bar{R}^{(2)} = \bar{R}^{(1)} + v^{(i-1)} \cdot \bar{R}^{(1)}\, w^{(i-1)}$$

... ...

$$\bar{R}^{(i)} = \bar{R}^{(i-1)} + v^{(1)} \cdot \bar{R}^{(i-1)}\, w^{(1)}$$

Step two : Use the factored $\bar{K}$ to solve the eqn.

$$\bar{K}\, \Delta \bar{d}^{(0)} = \bar{R}^{(i)}$$

Step three : $\Delta \bar{d}^{(1)} = \Delta \bar{d}^{(0)} + w^{(1)} \cdot \Delta \bar{d}^{(0)}\, v^{(1)}$

$$\Delta \bar{d}^{(2)} = \Delta \bar{d}^{(1)} + w^{(2)} \cdot \Delta \bar{d}^{(1)}\, v^{(2)}$$

... ...

$$\Delta \bar{d}^{(i)} = \Delta \bar{d}^{(i-1)} + w^{(i)} \cdot \Delta \bar{d}^{(i-1)}\, v^{(i)}$$

$\Delta d^{(i+1)} = \Delta \bar{d}^{(i)}$.

Remark 3: Memory cost: factored matrix $\bar{K}$

$$+ \quad \{ v^{(i)}, w^{(i)} \}_{i=1}^{i_{max}}$$

if i_{max} is reached, $\bar{K}$ should be reformed and refactored.

Remark 4: There are alternate options for the Choices of $v^{(i)}$ & $w^{(i)}$.

There are alternate options ~~to~~ other than BFGS that satisfy the Quasi-Newton equation.