Nonlinear Systems of Equations

\[ \overrightarrow{F}(\overrightarrow{x}) = \overrightarrow{0} \quad \text{where} \quad F : \mathbb{R}^n \rightarrow \mathbb{R}^n \]

for any given \( \overrightarrow{x}_k \), \( \overrightarrow{F}(\overrightarrow{x}) \) is called the residual of \( \overrightarrow{x}_k \).

\[ \overrightarrow{J}_F(\overrightarrow{x}) = \left( \frac{\partial f_i}{\partial x_j} \right) \in \mathbb{R}^{n \times n} \]

Newton's Method

\[ \overrightarrow{x}_{k+1} = \overrightarrow{x}_k - \overrightarrow{J}_F^{-1}(\overrightarrow{x}_k) \overrightarrow{F}(\overrightarrow{x}_k) \]

\( \overrightarrow{S}_k = -\overrightarrow{J}_F^{-1}(\overrightarrow{x}_k) \overrightarrow{F}(\overrightarrow{x}_k) \) is called the Newton direction, solution of

\[ \overrightarrow{J}_F(\overrightarrow{x}_k) \overrightarrow{S}_k = -\overrightarrow{F}(\overrightarrow{x}_k) \]

which is the most computationally expensive part of the algorithm.

We also choose an appropriate step size \( \lambda \in (0, 1] \) s.t.

\[ \overrightarrow{x}_{k+1} = \overrightarrow{x}_k + \lambda \overrightarrow{S}_k \]

Check if \( \| \overrightarrow{F}(\overrightarrow{x}_k) \| \) is sufficiently small.
STANDARD ASSUMPTIONS

1. \( F(x) = 0 \) has a solution \( x^* \).

2. \( F(x) \) is Lipschitz continuous in a neighborhood of \( x^* \).

3. \( F(x^*) \) is nonsingular.

Notice \( F \) has Lipschitz in induced-k norm of \( F \) since \( \| F(x) - F(x_0) \| \leq C \| x - x_0 \| \)
that is \( \| F \| = \sup \frac{\| Fx \|}{\| x \|} \) which is the largest singular value.

Letting the standard assumptions hold, if \( x_0 \) sufficiently close to \( x^* \),
then the Newton sequence exists, \( \{ x_k \} \) and converges to \( x^* \). In addition, there is \( k > 0 \) s.t.
\( \| e_{k+1} \| \leq k \| e_k \|^2 \) \( k \) suff. large.

Newton-Kantorovich Theorem - a stronger version of \( k \) theorem.

\( \frac{1}{1} \frac{1}{1} \frac{1}{1} \frac{1}{1} \)

PRACTICAL NEWTON'S

\( \text{Then, let the standard assumptions hold, let } \delta F_k \text{ be a perturbation of } F(x_k) \) and \( \delta F \)
a perturbation of \( F(x_k) \) such that \( \| \delta F \| < \delta \) and \( \| \delta F_k \| < \delta F \). Then
\( \text{if } x_0 \text{ is sufficiently close to } x^* \) and \( \delta F \) sufficiently small, then
\( x_{k+1} = x_k - (F(x_k) + \delta F)^{-1}(F(x_k) + \delta F) \)

is well-defined, and satisfies
\( \| e_{k+1} \| \leq C(\| e_k \|^2 + \| \delta F \| \| e_k \| + \| \delta F \|) \)
It is necessary that \( \lim_{k \to \infty} \| \delta F_k \| = 0 \) for convergence of \( x_k \).

If we have \( \| \delta J_f \| = O(1) \) we have linear convergence at best (\( C/\| \delta J \| < 1 \) is needed).

If \( \| \delta F \| = O(\| e_k \|^2) \) and \( \| \delta J \| = O(\| e_k \|) \) then we have quadratic convergence.

**Proof.**

\[
F(x^*) = F(x_k) + J_f(x_k)(x^* - x_k) + O(\| x^* - x_k \|^2) = 0
\]

\[
\delta F_k + \delta J(x^* - x_k) = F(x_k) + \delta F_k + (J_f(x_k) + \delta J)(x^* - x_k) + O(\| x^* - x_k \|^2)
\]

Thus

\[
(J_f(x_k) + \delta J)^{-1} [\delta F_k + \delta J(x^* - x_k)] = (J_f(x_k) + \delta J)^{-1} [F(x_k) + \delta F] + (x^* - x_k) + O(\| e_k \|^2)
\]

Let \( \delta J \) small enough that \( (J_f(x_k) + \delta J)^{-1} \) remains bounded.

Hence

\[
x^* = x_k - (J_f(x_k) + \delta J)^{-1} [F(x_k) + \delta F_k] + (J_f(x_k) + \delta J)^{-1} [\delta F_k + \delta J(x^* - x_k)] + O(\| e_k \|^2)
\]

\[
x_{k+1} = x_k - (J_f(x_k) + \delta J)^{-1} [F(x_k) + \delta F_k] + O(\| e_k \|^2)
\]

So

\[
\| e_{k+1} \| = O \left( (J_f(x_k) + \delta J)^{-1} \right)
\]

**Warning:** \( (J_f(x_k) + \delta J)^{-1} \) bounded

\[
\| e_{k+1} \| \leq C_j \left( \| \delta F_k \| + \| \delta J \| \| e_k \| + C \| e_k \|^2 \right)
\]