Termination of the Iteration

In general, we let the following criterion be the stopping condition

\[ \| F(x_k) \| \leq \tau_r \| F(x_0) \| + \tau_a \]

a relative tolerance \( \tau_r \) and an absolute tolerance \( \tau_a \)

Consider

\[ F(x_0) = F(x^*) + JF(x^*)(x_0 - x^*) + O(x_0 - x^*)^2 \]

\[ F(x_k) = JF(x^*)(x_k - x^*) + O(x_k - x^*)^2 \]

Thus

\[ \frac{\| F(x_k) \|}{\| F(x_0) \|} \approx \frac{\| JF(x^*) e_k \|}{\| JF(x^*) e_0 \|} \leq \frac{\| JF(x^*) \|^{-1} \| e_k \|}{\| e_0 \|} \]

and also

\[ \frac{\| F(x_k) \|}{\| F(x_0) \|} \geq \frac{1}{\| JF(x^*) \| JF(x^*)^{-1} \| e_k \|}{\| e_0 \|} \]

Call \( \kappa(F(x^*)) = \| JF(x^*) \| JF(x^*)^{-1} \| e_0 \| \) the condition number of the problem. (e.g., if e.g., 1000)

Assuming the problem is not ill-conditioned, \( \frac{\| F(x_k) \|}{\| F(x_0) \|} \) is a reasonable estimate of \( \frac{\| e_k \|}{\| e_0 \|} \)
Assume that Newton's method converges superlinearly.

\[ x_{k+1} = x_k + S_k = x_k - \frac{f(x_k)}{\nabla f(x_k)} \quad \text{o}(\|e_k\|) \quad \text{i.e.,} \lim_{k \to \infty} \frac{\|e_{k+1}\|}{\|e_k\|} = 0 \]

\[ x_{k+1} - x^* = x_k - x^* + \text{Newton step} \quad x_{k+1} - x_k \]

\[ b_k = -e_k + o(\|e_k\|) \]

\[ \rho_k = \frac{\|S_k\|}{\|S_k^+\|} \times \frac{\|e_{k+1}\|}{\|e_k\|} \geq \frac{\|e_{k+1}\|}{\|e_k\|} \]

Therefore, \[ \|e_{k+1}\| \leq \rho_k \|e_k\| \approx \frac{\|S_k\|}{\|S_{k+1}\|} \times \frac{\|S_{k+1}\|}{\|S_{k}^+\|} \]

Thus, can terminate Newton's iteration if \( \frac{\|S_k\|^2}{\|S_{k+1}\|} \) is sufficiently small.

**Armijo's Rule (line search)**

Calculate the step length separately from the direction.

Enhance robustness of convergence. Find the smallest \( m \) such that this condition holds

\[ \|F(x_k + 2^{-m}S_k)\| < (1 - \alpha 2^{-m}) \|F(x_k)\| \]

\( 0 < 1 < \alpha \). In practice, \( \alpha = 10^{-4} \) convenient.
A refined approach

Try step size 1. If no decrease in the residual, try step size \( \frac{1}{2} \), \( \frac{1}{4} \), etc...

If still not successful, calculate \( \lambda_{m+1} \), step size via

\[
\text{let } \phi(\lambda) = \| F(x_k + \lambda s_k) \|_2^2
\]

Use \( \phi(0), \phi(\lambda_m), \phi(\lambda_{m-1}) \) the most two recent step sizes

Then find the minimum of \( \phi(\lambda) \), which interpolates...