Additional Implementation Notes

for the PR Conjugate Gradient Method

  ① Always employ the Strong Wolfe conditions.

  ② Use $\quad 0 < c_1 < c_2 < \frac{1}{2} \quad$ typical: $c_1 = 10^{-4}$, $c_2 = \frac{2}{5}$.

  ③ Implement $\beta \leftarrow \max\{\beta, 0\}$ at each iteration.

  ④ Use a restart condition such as eq. 5.52.

For quasi-Newton (BFGS) Methods

  ① Always employ the strong Wolfe conditions.

  ② If $y_k^T s_k \leq 0$ then employ some restart strategy

    (a) Simply do not update H for the current iteration

    (b) Reset H_k to some H_0.

How can we choose typical values for x?

- The user may know reasonable values — in that case use them.

- Usually the initial iterate x_0 is composed of typical values.

- Typical values can be updated as the algorithm proceeds, but this is not usually important. Typical values can be very approximate — within one or two orders of magnitude.