To actually implement Newton's iteration (8.12), we need to solve an $n \times n$ linear system with the matrix $\boldsymbol{J_F}(\boldsymbol{x}^{(p)})$ for every $p = 0, 1, 2, \ldots$. To do so, we can employ the Gaussian elimination or some other method (iterative). However, for large dimensions, this may prove computationally expensive, especially if the Jacobian $\boldsymbol{J_F}(\boldsymbol{x}^{(p)})$ is ill-conditioned. To reduce the implementation cost of the original Newton's method, several modified versions of the algorithm have been proposed.

8.3.3 Modified Newton's Methods

Instead of inverting the Jacobian $\boldsymbol{J_F}(\boldsymbol{x}^{(p)})$ on every iteration, it may only be inverted once per several iterations. In the scalar context this means that we do not compute the new tangent to the graph of $F(x)$ for every $x^{(p)}$, $p = 0, 1, 2, \ldots$, as in the original method, see Figure 8.4, but rather "freeze" the same slope of the straight line for several iterations, and after that update. In the context of systems, this approach may considerably reduce the computational cost. Its downside, though, is a slower convergence of the method, which is the case for both scalar equations and systems.

Alternatively, when the linear system:

$$\boldsymbol{F}(\boldsymbol{x}^{(p)}) + \boldsymbol{J_F}(\boldsymbol{x}^{(p)})(\boldsymbol{x} - \boldsymbol{x}^{(p)}) = \boldsymbol{0}$$

is solved by iterations, instead of iterating until, say, a predetermined level of the residual reduction is reached, we can only run a fixed number of linear iterations chosen ahead of time. While reducing the cost, this, of course, only yields a fairly "crude" approximation of the solution to the linear system, and also slows down the overall convergence. Depending on what particular linear iteration is used (Chapter 6), there are Newton-Jacobi, Newton-SOR, Newton-Krylov, and other methods of this type.

In addition to the costs associated with inverting the Jacobian $\boldsymbol{J_F}(\boldsymbol{x}^{(p)})$, the cost of computing the matrix $\boldsymbol{J_F}(\boldsymbol{x}^{(p)})$ itself for every $p = 0, 1, 2, \ldots$ may be quite noticeable. To reduce it, instead of evaluating the Jacobian exactly, one often evaluates it approximately with the help of finite differences. Either forward differencing or central differencing can be employed; the analysis of the corresponding errors can be found in Chapter 9. To keep the quadratic convergence of the method when the Jacobian is evaluated using finite differences, one needs to satisfy certain constraints for the step size, otherwise the convergence rate drops to linear.

Finally, there is a family of the so-called quasi-Newton methods that can be built for both exact and finite-difference Jacobians [cf. formula (8.12)]:

$$\boldsymbol{x}^{(p+1)} = \boldsymbol{x}^{(p)} - \gamma_p \boldsymbol{J_F}^{-1}(\boldsymbol{x}^{(p)}) \boldsymbol{F}(\boldsymbol{x}^{(p)}), \quad p = 0, 1, 2, \ldots.$$

The quantities γ_p are known as the damping parameters.

Besides Newton's method, other methods are available for the solution of nonlinear systems. There is, for example, a multidimensional version of the secant method called the Broyden method. We do not discuss it in this book. We rather refer the reader to the specialized literature for further detail on the numerical solution of the nonlinear scalar equations and systems, see, e.g., the monographs [OR00, Kel95, Kel03].

Exercises

1.* Prove quadratic convergence of Newton's method for systems, i.e., establish estimate (8.13) given the conditions right after equation (8.13) on page 245 and also assuming that the initial guess $\boldsymbol{x}^{(0)}$ is sufficiently close to the solution $\hat{\boldsymbol{x}}$.

   Hint. Increment of $F(x)$ is to be obtained by integrating the corresponding directional derivative.

2. Build an algorithm for computing $\sqrt{5}$ with a given precision. Interpret $\sqrt{5}$ as a solution to the equation $x^2-5=0$ and employ Newton's method. Show that the iterations will converge for any initial guess $x^{(0)}>0$.

3. Let $\phi=\phi(x)$ and $\psi=\psi(x)$ be two functions with bounded second derivatives. For solving the equation $F(x)\equiv\phi(x)-\psi(x)=0$ one can use Newton's method.

   a) Let the graphs of the functions $\phi=\phi(x)$ and $\psi=\psi(x)$ be plotted on the Cartesian plane (x,y), and let their intersection point [the root of $F(x)=0$] have the abscissa $x=\hat{x}$. Assume that the Newton iterate $x^{(p)}$ is already computed and provide a geometric interpretation of how one obtains the next iterate $x^{(p+1)}$. Use the following form of Newton's linearization:

   $$\phi(x^{(p)})-\psi(x^{(p)})+(\phi'(x^{(p)})-\psi'(x^{(p)}))(x^{(p+1)}-x^{(p)})=0.$$

   b) Assume, for definiteness, that $\phi(x)>\psi(x)$ for $x>\hat{x}$, and $\phi'(\hat{x})-\psi'(\hat{x})>0$. Let also the graph of $\phi(x)$ be convex, i.e., $\phi''(x)>0$, and the graph of $\psi(x)$ be concave, i.e., $\psi''(x)<0$. Show that Newton's method will converge to $\hat{x}$ for any initial guess $x^{(0)}>\hat{x}$.

4. Use Newton's method to compute real solutions to the following systems of nonlinear equations. Obtain the results with five significant digits:

   a) $\sin x-y=1.30$; $\cos y-x=-0.84$.

   b) $x^2+4y^2=1$; $x^4+y^4=0.5$.

5. A nonlinear boundary value problem:

   $$\frac{d^2u}{dx^2}-u^3=x^2,\quad x\in[0,1],$$
   $$u(0)=1,\ \ u(1)=3,$$

   with the unknown function $u=u(x)$ is approximated on the uniform grid $x_j=j\cdot h$, $j=0,1,\ldots,N$, $h=\frac{1}{N}$, using the following second-order central-difference scheme:

   $$\frac{u_{j+1}-2u_j+u_{j-1}}{h^2}-u_j^3=(jh)^2,\ \ j=1,2,\ldots,N-1,$$
   $$u_0=1,\ \ u_N=3,$$

   where u_j, $j=0,1,2,\ldots,N$, is the unknown discrete solution on the grid.

   Solve the foregoing discrete system of nonlinear equations with respect to u_j, $j=0,1,2,\ldots,N$, using Newton's method. Introduce the following initial guess: $u_j^{(0)}=1+2(jh)$, $j=0,1,2,\ldots,N$. This grid function obviously satisfies the boundary conditions: