Optimization Techniques
Lecture 5&6
Hale Gonce Köçken

Hale Gonce Köçken

The Newton's method

The conjugate direction methods

Newton's Method

Recall that the method of steepest descent uses only first derivatives (gradients) in selecting a suitable search direction. This strategy is not always the most effective. If higher derivatives are used, the resulting iterative algorithm may perform better than the steepest descent method.

Newton's method (sometimes called the Newton-Raphson method) uses first and second derivatives and indeed does perform better than the steepest descent method if the initial point is close to the minimizer.

idea behind this method is as follows. Given a starting point, we construct a quadratic approximation to the objective function that matches the first and second derivative values at that point.

We then minimize the approximate (quadratic) function instead of the original objective function. We use the minimizer of the approximate function as the starting point in the next step and repeat the procedure iteratively.

If the objective function is quadratic, then the approximation is exact, and the method yields the true minimizer in one step.

If, on the other hand, the objective function is not quadratic, then the approximation will provide only an estimate of the position of the true minimizer.

Figure 9.1 illustrates the above idea.

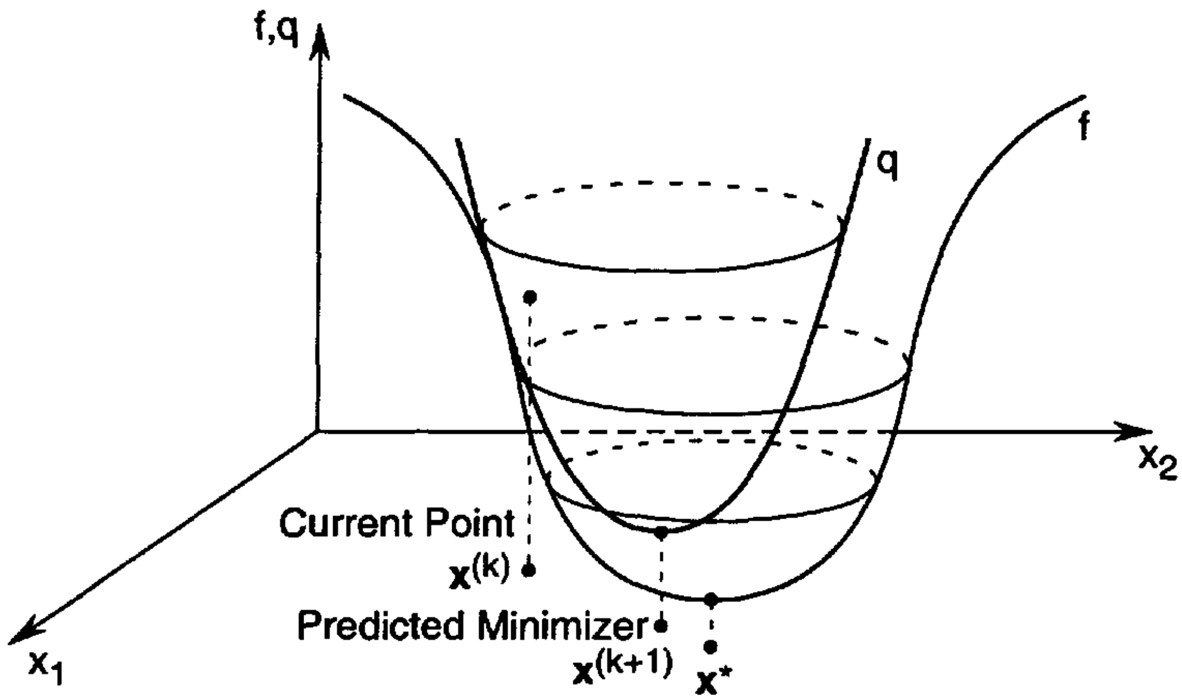

Figure 9.1 Quadratic approximation to the objective function using first and second derivatives

We can obtain a quadratic approximation to the given twice continuously differentiable objection function $f : \mathbb{R}^n \to \mathbb{R}$ using the Taylor series expansion of f about the current point $\boldsymbol{x}^{(k)}$, neglecting terms of order three and higher. We obtain

$$f(\boldsymbol{x}) \approx f(\boldsymbol{x}^{(k)}) + (\boldsymbol{x} - \boldsymbol{x}^{(k)})^T \boldsymbol{g}^{(k)} + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^{(k)})^T \boldsymbol{F}(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}) \triangleq q(\boldsymbol{x}),$$

where, for simplicity, we use the notation $\boldsymbol{g}^{(k)} = \nabla f(\boldsymbol{x}^{(k)})$. Applying the FONC to q yields

$$\boldsymbol{0} = \nabla q(\boldsymbol{x}) = \boldsymbol{g}^{(k)} + \boldsymbol{F}(\boldsymbol{x}^{(k)})(\boldsymbol{x} - \boldsymbol{x}^{(k)}).$$

If $\boldsymbol{F}(\boldsymbol{x}^{(k)}) > 0$, then q achieves a minimum at

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \boldsymbol{F}(\boldsymbol{x}^{(k)})^{-1} \boldsymbol{g}^{(k)}.$$

This recursive formula represents Newton's method.

Example: Find the minimizer of $f(x) = 2x_1^2 + 2x_1x_2 + 2x_2^2 - 4x_1 - 6x_2$ with the starting point $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ by using Newton's method.

Conjugate Direction Methods

The class of *conjugate direction methods* can be viewed as being intermediate between the method of steepest descent and Newton's method. The conjugate direction methods have the following properties:

1. Solve quadratics of n variables in n steps;
2. The usual implementation, the *conjugate gradient algorithm*, requires no Hessian matrix evaluations;
3. No matrix inversion and no storage of an $n \times n$ matrix required.

Basically, two directions $\boldsymbol{d}^{(1)}$ and $\boldsymbol{d}^{(2)}$ in $\mathbb{R}^n$ are said to be $\boldsymbol{Q}$-conjugate if $\boldsymbol{d}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(2)} = 0$.

Definition 10.1 Let $\boldsymbol{Q}$ be a real symmetric $n \times n$ matrix. The directions $\boldsymbol{d}^{(0)}, \boldsymbol{d}^{(1)}, \boldsymbol{d}^{(2)}, \ldots, \boldsymbol{d}^{(m)}$ are *$\boldsymbol{Q}$-conjugate* if, for all $i \neq j$, we have $\boldsymbol{d}^{(i)T}\boldsymbol{Q}\boldsymbol{d}^{(j)} = 0$.

Lemma 10.1 *Let $\boldsymbol{Q}$ be a symmetric positive definite $n \times n$ matrix. If the directions $\boldsymbol{d}^{(0)}, \boldsymbol{d}^{(1)}, \ldots, \boldsymbol{d}^{(k)} \in \mathbb{R}^n$, $k \leq n-1$, are nonzero and $\boldsymbol{Q}$-conjugate, then they are linearly independent.*

Proof. Let $\alpha_0, \ldots, \alpha_k$ be scalars such that

$$\alpha_0 \boldsymbol{d}^{(0)} + \alpha_1 \boldsymbol{d}^{(1)} + \ldots + \alpha_k \boldsymbol{d}^{(k)} = \boldsymbol{0}.$$

Premultiplying the above equality by $\boldsymbol{d}^{(j)T}\boldsymbol{Q}$, $0 \leq j \leq k$, yields

$$\alpha_j \boldsymbol{d}^{(j)T} \boldsymbol{Q} \boldsymbol{d}^{(j)} = 0,$$

because all other terms $\boldsymbol{d}^{(j)T}\boldsymbol{Q}\boldsymbol{d}^{(i)} = 0$, $i \neq j$, by $\boldsymbol{Q}$-conjugacy. But $\boldsymbol{Q} = \boldsymbol{Q}^T > 0$ and $\boldsymbol{d}^{(j)} \neq \boldsymbol{0}$; hence $\alpha_j = 0$, $j = 0, 1, \ldots, k$. Therefore, $\boldsymbol{d}^{(0)}, \boldsymbol{d}^{(1)}, \ldots, \boldsymbol{d}^{(k)}$, $k \leq n-1$, are linearly independent.

Example 10.1 Let

$$Q = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Note that $Q = Q^T > 0$. The matrix Q is positive definite because all its leading principal minors are positive:

Our goal is to construct a set of Q-conjugate vectors $d^{(0)}, d^{(1)}, d^{(2)}$.

Let $d^{(0)} = [1, 0, 0]^T$, $d^{(1)} = [d_1^{(1)}, d_2^{(1)}, d_3^{(1)}]^T$, $d^{(2)} = [d_1^{(2)}, d_2^{(2)}, d_3^{(2)}]^T$. We require $d^{(0)T} Q d^{(1)} = 0$. We have

$$d^{(0)T} Q d^{(1)} = [1, 0, 0] \begin{bmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} d_1^{(1)} \\ d_2^{(1)} \\ d_3^{(1)} \end{bmatrix}$$

Let $d_1^{(1)} = 1, d_2^{(1)} = 0, d_3^{(1)} = -3$. Then, $\boldsymbol{d}^{(1)} = [1, 0, -3]^T$, and thus $\boldsymbol{d}^{(0)T}\boldsymbol{Q}\boldsymbol{d}^{(1)} = 0$.

To find the third vector $\boldsymbol{d}^{(2)}$, which would be $\boldsymbol{Q}$-conjugate with $\boldsymbol{d}^{(0)}$ and $\boldsymbol{d}^{(1)}$, we require $\boldsymbol{d}^{(0)T}\boldsymbol{Q}\boldsymbol{d}^{(2)} = 0$ and $\boldsymbol{d}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(2)} = 0$. We have

$$\begin{aligned} \boldsymbol{d}^{(0)T}\boldsymbol{Q}\boldsymbol{d}^{(2)} &= 3d_1^{(2)} + d_3^{(2)} = 0, \\ \boldsymbol{d}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(2)} &= -6d_2^{(2)} - 8d_3^{(2)} = 0. \end{aligned}$$

If we take $\boldsymbol{d}^{(2)} = [1, 4, -3]^T$, then the resulting set of vectors is mutually conjugate.

A systematic procedure for finding $\boldsymbol{Q}$-conjugate vectors

$$\boldsymbol{Q}\boldsymbol{x} = \boldsymbol{b}$$

1. Basic Conjugate Direction Algorithm

We now present the conjugate direction algorithm for minimizing the quadratic function of n variables

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{b},$$

where $\boldsymbol{Q} = \boldsymbol{Q}^T > 0, \boldsymbol{x} \in \mathbb{R}^n$. Note that because $\boldsymbol{Q} > 0$,

Given a starting point $\boldsymbol{x}^{(0)}$, and $\boldsymbol{Q}$-conjugate directions $\boldsymbol{d}^{(0)}, \boldsymbol{d}^{(1)}, \ldots, \boldsymbol{d}^{(n-1)}$; for $k \geq 0$,

$$\begin{aligned} \boldsymbol{g}^{(k)} &= \nabla f(\boldsymbol{x}^{(k)}) = \boldsymbol{Q}\boldsymbol{x}^{(k)} - \boldsymbol{b}, \\ \alpha_k &= -\frac{\boldsymbol{g}^{(k)T}\boldsymbol{d}^{(k)}}{\boldsymbol{d}^{(k)T}\boldsymbol{Q}\boldsymbol{d}^{(k)}}, \\ \boldsymbol{x}^{(k+1)} &= \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}^{(k)}. \end{aligned}$$

Theorem 10.1 *For any starting point $\boldsymbol{x}^{(0)}$, the basic conjugate direction algorithm converges to the unique $\boldsymbol{x}^*$ (that solves $\boldsymbol{Q}\boldsymbol{x} = \boldsymbol{b}$) in n steps; that is, $\boldsymbol{x}^{(n)} = \boldsymbol{x}^*$.*

Example 10.2 Find the minimizer of

$$f(x_1, x_2) = \frac{1}{2}\boldsymbol{x}^T \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \boldsymbol{x} - \boldsymbol{x}^T \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \boldsymbol{x} \in \mathbb{R}^2,$$

using the conjugate direction method with the initial point $\boldsymbol{x}^{(0)} = [0,0]^T$, and $\boldsymbol{Q}$-conjugate directions $\boldsymbol{d}^{(0)} = [1,0]^T$ and $\boldsymbol{d}^{(1)} = \left[-\frac{3}{8}, \frac{3}{4}\right]^T$.

We have $\boldsymbol{g}^{(0)} = -\boldsymbol{b} = [1, -1]^T$,

and hence $\alpha_0 = -\dfrac{\boldsymbol{g}^{(0)T}\boldsymbol{d}^{(0)}}{\boldsymbol{d}^{(0)T}\boldsymbol{Q}\boldsymbol{d}^{(0)}} =$

Thus, $\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha_0 \boldsymbol{d}^{(0)} =$

To find $\boldsymbol{x}^{(2)}$, we compute $\boldsymbol{g}^{(1)} = \boldsymbol{Q}\boldsymbol{x}^{(1)} - \boldsymbol{b} =$

and $\alpha_1 = -\frac{\boldsymbol{g}^{(1)T}\boldsymbol{d}^{(1)}}{\boldsymbol{d}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(1)}} =$

Therefore, $\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} + \alpha_1 \boldsymbol{d}^{(1)} =$

Because f is a quadratic function in two variables, $\boldsymbol{x}^{(2)} = \boldsymbol{x}^{*}$.

2. Conjugate Gradient Algorithm

Our first search direction from an initial point $\boldsymbol{x}^{(0)}$ is in the direction of steepest descent; that is,

$$\boldsymbol{d}^{(0)} = -\boldsymbol{g}^{(0)}.$$

Thus, $\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha_0 \boldsymbol{d}^{(0)}$, where

$$\alpha_0 = \arg\min_{\alpha \geq 0} f(\boldsymbol{x}^{(0)} + \alpha \boldsymbol{d}^{(0)}) = -\frac{\boldsymbol{g}^{(0)T}\boldsymbol{d}^{(0)}}{\boldsymbol{d}^{(0)T}\boldsymbol{Q}\boldsymbol{d}^{(0)}}.$$

In the next stage, we search in a direction $\boldsymbol{d}^{(1)}$ that is $\boldsymbol{Q}$-conjugate to $\boldsymbol{d}^{(0)}$. We choose $\boldsymbol{d}^{(1)}$ as a linear combination of $\boldsymbol{g}^{(1)}$ and $\boldsymbol{d}^{(0)}$. In general, at the $(k+1)$st step, we choose $\boldsymbol{d}^{(k+1)}$ to be a linear combination of $\boldsymbol{g}^{(k+1)}$ and $\boldsymbol{d}^{(k)}$. Specifically, we choose

$$\boldsymbol{d}^{(k+1)} = -\boldsymbol{g}^{(k+1)} + \beta_k \boldsymbol{d}^{(k)}, \quad k = 0, 1, 2, \ldots.$$

The coefficients β_k, $k = 1, 2, \ldots$, are chosen in such a way that $\boldsymbol{d}^{(k+1)}$ is $\boldsymbol{Q}$-conjugate to $\boldsymbol{d}^{(0)}, \boldsymbol{d}^{(1)}, \ldots, \boldsymbol{d}^{(k)}$. This is accomplished by choosing β_k to be

$$\beta_k = \frac{\boldsymbol{g}^{(k+1)T}\boldsymbol{Q}\boldsymbol{d}^{(k)}}{\boldsymbol{d}^{(k)T}\boldsymbol{Q}\boldsymbol{d}^{(k)}}.$$

The conjugate gradient algorithm is summarized below.

1. Set $k := 0$; select the initial point $\boldsymbol{x}^{(0)}$.
2. $\boldsymbol{g}^{(0)} = \nabla f(\boldsymbol{x}^{(0)})$. If $\boldsymbol{g}^{(0)} = \boldsymbol{0}$, stop, else set $\boldsymbol{d}^{(0)} = -\boldsymbol{g}^{(0)}$.
3. $\alpha_k = -\dfrac{\boldsymbol{g}^{(k)T}\boldsymbol{d}^{(k)}}{\boldsymbol{d}^{(k)T}\boldsymbol{Q}\boldsymbol{d}^{(k)}}$.
4. $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}^{(k)}$.
5. $\boldsymbol{g}^{(k+1)} = \nabla f(\boldsymbol{x}^{(k+1)})$. If $\boldsymbol{g}^{(k+1)} = \boldsymbol{0}$, stop.
6. $\beta_k = \dfrac{\boldsymbol{g}^{(k+1)T}\boldsymbol{Q}\boldsymbol{d}^{(k)}}{\boldsymbol{d}^{(k)T}\boldsymbol{Q}\boldsymbol{d}^{(k)}}$.
7. $\boldsymbol{d}^{(k+1)} = -\boldsymbol{g}^{(k+1)} + \beta_k \boldsymbol{d}^{(k)}$.
8. Set $k := k + 1$; go to step 3.

Example 10.3 Consider the quadratic function

$$f(x_1, x_2, x_3) = \frac{3}{2}x_1^2 + 2x_2^2 + \frac{3}{2}x_3^2 + x_1x_3 + 2x_2x_3 - 3x_1 - x_3.$$

We find the minimizer using the conjugate gradient algorithm, using the starting point $\boldsymbol{x}^{(0)} = [0, 0, 0]^T$.

We can represent f as $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{b}$,

where $\boldsymbol{Q} = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$

We have

$$\boldsymbol{g}(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) = \boldsymbol{Q}\boldsymbol{x} - \boldsymbol{b} = [3x_1 + x_3 - 3, 4x_2 + 2x_3, x_1 + 2x_2 + 3x_3 - 1]^T.$$

Hence, $\boldsymbol{g}^{(0)} = [-3, 0, -1]^T, \qquad \boldsymbol{d}^{(0)} = -\boldsymbol{g}^{(0)},$

$$\alpha_0 = -\frac{\boldsymbol{g}^{(0)T}\boldsymbol{d}^{(0)}}{\boldsymbol{d}^{(0)T}\boldsymbol{Q}\boldsymbol{d}^{(0)}} = \frac{10}{36} = 0.2778,$$

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha_0\boldsymbol{d}^{(0)} = [0.8333, 0, 0.2778]^T.$$

The next stage yields

$$\boldsymbol{g}^{(1)} = \nabla f(\boldsymbol{x}^{(1)}) = [-0.2222, 0.5556, 0.6667]^T,$$

$$\beta_0 = \frac{\boldsymbol{g}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(0)}}{\boldsymbol{d}^{(0)T}\boldsymbol{Q}\boldsymbol{d}^{(0)}} = 0.08025.$$

We can now compute

$$\boldsymbol{d}^{(1)} = -\boldsymbol{g}^{(1)} + \beta_0\boldsymbol{d}^{(0)} = [0.4630, -0.5556, -0.5864]^T.$$

Hence,

$$\alpha_1 = -\frac{\boldsymbol{g}^{(1)T}\boldsymbol{d}^{(1)}}{\boldsymbol{d}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(1)}} = 0.2187,$$

$$\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(1)} + \alpha_1\boldsymbol{d}^{(1)} = [0.9346, -0.1215, 0.1495]^T.$$

To perform the third iteration, we compute

$$\begin{aligned}
\boldsymbol{g}^{(2)} &= \nabla f(\boldsymbol{x}^{(2)}) = [-0.04673, -0.1869, 0.1402]^T, \\
\beta_1 &= \frac{\boldsymbol{g}^{(2)T}\boldsymbol{Q}\boldsymbol{d}^{(1)}}{\boldsymbol{d}^{(1)T}\boldsymbol{Q}\boldsymbol{d}^{(1)}} = 0.07075, \\
\boldsymbol{d}^{(2)} &= -\boldsymbol{g}^{(2)} + \beta_1 \boldsymbol{d}^{(1)} = [0.07948, 0.1476, -0.1817]^T.
\end{aligned}$$

Hence,

$$\alpha_2 = -\frac{\boldsymbol{g}^{(2)T}\boldsymbol{d}^{(2)}}{\boldsymbol{d}^{(2)T}\boldsymbol{Q}\boldsymbol{d}^{(2)}} = 0.8231,$$

and

$$\boldsymbol{x}^{(3)} = \boldsymbol{x}^{(2)} + \alpha_2 \boldsymbol{d}^{(2)} = [1.000, 0.000, 0.000]^T.$$

Note that

$$\boldsymbol{g}^{(3)} = \nabla f(\boldsymbol{x}^{(3)}) = \boldsymbol{0},$$

as expected, because f is a quadratic function of three variables. Hence, $\boldsymbol{x}^* = \boldsymbol{x}^{(3)}$.

3. Conjugate Gradient Algorithm for non-quadratic problems

For a quadratic, the matrix Q, the Hessian of the quadratic, is constant.

However, for a general nonlinear function the Hessian is a matrix that has to be reevaluated at each iteration of the algorithm.

This can be computationally very expensive. Thus, an efficient implementation of the conjugate gradient algorithm that eliminates the Hessian evaluation at each step is desirable.

The Hestenes-Stiefel formula.

Recall that $\beta_k = \frac{\boldsymbol{g}^{(k+1)T} \boldsymbol{Q} \boldsymbol{d}^{(k)}}{\boldsymbol{d}^{(k)T} \boldsymbol{Q} \boldsymbol{d}^{(k)}}$.

$$\boldsymbol{Q}\boldsymbol{d}^{(k)} \longleftrightarrow (\boldsymbol{g}^{(k+1)} - \boldsymbol{g}^{(k)})/\alpha_k.$$

The two terms are equal in the quadratic case, as we now show. Now, $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}^{(k)}$. Premultiplying both sides by $\boldsymbol{Q}$, and recognizing that $\boldsymbol{g}^{(k)} = \boldsymbol{Q}\boldsymbol{x}^{(k)} - \boldsymbol{b}$, we get $\boldsymbol{g}^{(k+1)} = \boldsymbol{g}^{(k)} + \alpha_k \boldsymbol{Q}\boldsymbol{d}^{(k)}$, which we can rewrite as $\boldsymbol{Q}\boldsymbol{d}^{(k)} = (\boldsymbol{g}^{(k+1)} - \boldsymbol{g}^{(k)})/\alpha_k$. Substituting this into the original equation for β_k gives

$$\beta_k = \frac{\boldsymbol{g}^{(k+1)T}[\boldsymbol{g}^{(k+1)} - \boldsymbol{g}^{(k)}]}{\boldsymbol{d}^{(k)T}[\boldsymbol{g}^{(k+1)} - \boldsymbol{g}^{(k)}]},$$

which is called the Hestenes-Stiefel formula.

The Polak-Ribière formula.

$$\beta_k = \frac{\boldsymbol{g}^{(k+1)T}[\boldsymbol{g}^{(k+1)} - \boldsymbol{g}^{(k)}]}{\boldsymbol{g}^{(k)T}\boldsymbol{g}^{(k)}}.$$

The Fletcher-Reeves formula.

$$\beta_k = \frac{\boldsymbol{g}^{(k+1)T}\boldsymbol{g}^{(k+1)}}{\boldsymbol{g}^{(k)T}\boldsymbol{g}^{(k)}},$$

For nonquadratic problems, the algorithm will not usually converge in n steps, and as the algorithm progresses, the "$\boldsymbol{Q}$-conjugacy" of the direction vectors will tend to deteriorate. Thus, a common practice is to reinitialize the direction vector to the negative gradient after every few iterations (e.g., n or $n+1$), and continue until the algorithm satisfies the stopping criterion.

Example:

Let $f(\boldsymbol{x})$, $\boldsymbol{x} = [x_1, x_2]^T \in \mathbb{R}^2$, be given by

$$f(\boldsymbol{x}) = \frac{5}{2}x_1^2 + \frac{1}{2}x_2^2 + 2x_1x_2 - 3x_1 - x_2.$$

a. Express $f(\boldsymbol{x})$ in the form of $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{b}$.

b. Find the minimizer of f using the conjugate gradient algorithm. Use a starting point of $\boldsymbol{x}^{(0)} = [0, 0]^T$.

c. Calculate the minimizer of f analytically from $\boldsymbol{Q}$ and $\boldsymbol{b}$, and check it with your answer in part b.