Lecture-5

Quadratic Functions

Quadratic Functions

\[ f(x) = \frac{1}{2} x^T Q x - b^T x \]

\[ \nabla f(x) = Q x - b \]

If \( x^* \) is a unique solution of \( Q x = b \), then it is a stationary point of \( f \).

If the linear system \( Q x = b \) cannot be solved, then function does not have a stationary point, it is unbounded.
Quadratic Functions

\[ f(x) = \frac{1}{2} x^T Q x - b^T x \]

\[ \nabla f(x) = Q x - b \]

According to definition, for any vector \( x \) and \( p \):

\[ f(x + \alpha p) = \frac{1}{2} (x + \alpha p)^T Q (x + \alpha p) - b^T (x + \alpha p) \]

If \( x^* \) is stationary point

\[ f(x^* + \alpha p) = f(x^*) + \alpha p^T (Q x^* - b) + \frac{1}{2} \alpha^2 p^T Q p \]

\[ f(x^* + \alpha p) = f(x^*) + \frac{1}{2} \alpha^2 p^T Q p \]
Quadratic Functions

\[ f(x^* + \alpha p) = f(x^*) + \frac{1}{2} \alpha^2 p^T Qp \]

The behavior of \( f \) is determined by matrix \( Q \)

Let \( Qu_j = \lambda_j u_j \)

Let \( p \) is equal to \( u_j \)

\[ f(x^* + \alpha u_j) = f(x^*) + \frac{1}{2} \alpha^2 u_j^T Qu_j \]

\[ f(x^* + \alpha u_j) = f(x^*) + \frac{1}{2} \alpha^2 u_j^T \lambda_j u_j \]

\[ f(x^* + \alpha u_j) = f(x^*) + \frac{1}{2} \alpha^2 \lambda_j \]

\( Q \) is symmetric

---

Quadratic Functions

- The change in \( f \) when moving away from \( x^* \) along the direction \( u_j \) depends on the sign of \( \lambda_j \)
  - If \( \lambda_j \) is positive \( f \) will strictly increase as \( |\alpha| \) increases
  - If \( \lambda_j \) is negative, \( f \) is decreasing as \( |\alpha| \) increases.
  - If \( \lambda_j \) is zero, the value of \( f \) remains constant when moving along any direction parallel to \( u_j \)
  - \( f \) reduces to a linear function along any such direction, since quadratic term vanishes.

\[ f(x^* + \alpha u_j) = f(x^*) + \frac{1}{2} \alpha^2 \lambda_j \]
Quadratic Functions

- When all eigenvalues of $Q$ are positive, $x^*$ is the unique global minimum.
  - The contours of $f$ are ellipsoid whose principal axes are in the directions of the eigenvectors of $Q$, with lengths proportional to square root of corresponding eigenvalues.
- If $Q$ is positive semi-definite, a stationary point (if it exists) is a weak local minimum.
- If $Q$ is indefinite and non-singular, $x^*$ is a saddle point, $f$ is unbounded.
  \[ f(x^* + \alpha u_j) = f(x^*) + \frac{1}{2} \alpha^2 \lambda_j \]

Iso Contours (Contour Map)

\[ f(x_1, x_2) = c \]

\[ f(x_1, x_2) = e^{x_1}(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1) \]
\[ c = .2, .4, 1, 1.7, 1.8, 2, 3, 4, 5, 6, 20 \]
Quadratic Functions

Two positive eigenvalues

\[ Q = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} -5.5 \\ -3.5 \end{bmatrix} \]

Eigenvectors

\(-0.8507 \quad 0.5257\)
\(-0.5257 \quad -0.8507\)

Eigenvalue 6.8541, 0.1459

Semi PD

Fig. 21: Contours of (i) a positive definite quadratic function; (ii) a positive semi-definite quadratic function; and (iii) an indefinite quadratic function.

Quadratic Functions

One positive eigenvalue, one zero eigenvalue

\[ Q = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} -4 \\ -2 \end{bmatrix} \]

Eigenvectors

\(-0.8944 \quad 0.4472\)
\(-0.4472 \quad -0.8944\)

Eigenvalue 5, 0

Fig. 22: Contours of (i) a positive definite quadratic function; (ii) a positive semi-definite quadratic function; and (iii) an indefinite quadratic function.
Quadratic Functions

One positive eigenvalue, one zero eigenvalue

\[ Q = \begin{bmatrix} 3 & -1 \\ -1 & -8 \end{bmatrix}, \quad b = \begin{bmatrix} -.5 \\ 8.5 \end{bmatrix} \]

Indefinite

Eigenvalue 3.0902, -8.0902

Eigenvectors

-0.9960 -0.0898
0.0898 -0.9960

How about a function with \( Q \), which is a diagonal matrix?
Steepest Descent

How about a function with $Q$, which is a multiple of an identity matrix?
Steepest Descent

Convergence Rate of Steepest Descent

\[ f(x) = \frac{1}{2} x^T Qx - b^T x \]

\[ \nabla f(x) = Qx - b \]

\( x^* \) is a unique solution of \( Qx = b \)

Let us compute step length, which minimizes the function:

\[ f(x_k - \alpha g_k) = \frac{1}{2} (x_k - \alpha g_k)^T Q(x_k - \alpha g_k) - b^T (x_k - \alpha g_k) \]
Convergence Rate of Steepest Descent

\[
\frac{d}{d\alpha} f(x_k - \alpha g_k) = \frac{d}{d\alpha} \left( \frac{1}{2} (x_k - \alpha g_k)^T Q (x_k - \alpha g_k) - b^T (x_k - \alpha g_k) \right) = 0
\]

\[
= - (x_k - \alpha g_k)^T Q g_k + b^T g_k = 0
\]

\[
- x_k^T Q g_k + \alpha g_k^T Q g_k + b^T g_k = 0
\]

\[
\alpha g_k^T Q g_k = x_k^T Q g_k - b^T g_k
\]

\[
\alpha = \frac{x_k^T Q g_k - b^T g_k}{g_k^T Q g_k}
\]

\[
\alpha = \frac{(x_k^T Q - b^T) g_k}{g_k^T Q g_k}
\]

\[
\alpha = \frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T Q \nabla f_k}
\]

\[
x_{k+1} = x_k - \alpha_k \nabla f_k
\]

\[
x_{k+1} = x_k - \frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T Q \nabla f_k} \nabla f_k
\]

\[
\text{Define}
\]

\[
\frac{1}{2} \| x - x^* \|_Q^2 = f(x) - f(x^*)
\]

Using:

\[
x_{k+1} = x_k - \frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T Q \nabla f_k} \nabla f_k
\]

It can be shown (homework):

\[
\| x_{k+1} - x^* \|_Q^2 \leq \left( 1 - \frac{\left( \nabla f_k^T \nabla f_k \right)^2}{\left( \nabla f_k^T Q \nabla f_k \right) \left( \nabla f_k^T Q^{-1} \nabla f_k \right)} \right) \| x_k - x^* \|_Q^2
\]

OR

\[
\| x_{k+1} - x^* \|_Q^2 \leq \left( \frac{\lambda_{\min} - \lambda_{\max}}{\lambda_{\min} + \lambda_{\max}} \right)^2 \| x_k - x^* \|_Q^2
\]

where \( 0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \lambda_n \) are eigenvalues of \( Q \).
Convergence Rate of Steepest Descent

\[ \| x_{k+1} - x^* \|_Q^2 \leq \left( \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \right)^2 \| x_k - x^* \|_Q^2 \]

As the condition number increases the contours of the quadratic become more elongated, the zigzags of line search becomes more pronounced.

Theorem 3.4: Steepest Descent

\[ f(x_{k+1}) - f(x^*) \leq \left( \frac{\lambda_n - \lambda_1}{\lambda_n + \lambda_1} \right)^2 (f(x_k) - f(x^*)) \]

where \(0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \lambda_n\) are eigenvalues of Hessian

If the condition number is 800, and \(f(x) = 1\) and \(f(x^*) = 0\), After 1000 iterations the value of function will be .08.