Convexity and optimality

Global minimizer:
\( x^* \in S \) such that
\[ f(x^*) \leq f(x) \quad \forall x \in S \subset \mathbb{R}^n \]

Local minimizer:
\( x^* \in S \) such that, for some \( \epsilon > 0 \),
\[ f(x^*) \leq f(x) \quad \forall x \in S \cap B(x^*; \epsilon) \]

Theorem. For convex functions on convex sets holds that each local minimizer is a global minimizer.

Example:
\[ \min_x f(x) \text{ such that } Ax \leq b, \]

where \( f(x) = c^T x \) (linear program) or
\[ f(x) = \gamma + c^T x + \frac{1}{2} x^T Q x, \text{ } Q \text{ positive semidefinite} \] (quadratic program)
Iterative algorithms

Problem:

\[ \min_{x \in S} f(x) \]

Many optimization algorithms are of the type

1. Specify an initial guess \( x_0 \)
2. For \( k = 0, 1, \ldots \)
   2.1 If \( x_k \) optimal stop
   2.2 Determine a search direction \( p_k \) and a step \( \alpha_k \) and set
   2.3 \( x_{k+1} = x_k + \alpha_k p_k \)
Convergence rate

Definition
The sequence \( \{ x_k \} \) converges to \( x^* \) with rate \( p \) and rate constant \( C \) if
\[
\lim_{k \to \infty} \frac{\| x_{k+1} - x^* \|}{\| x_k - x^* \|^p} = C.
\]

- **Linear**: \( p = 1 \) and \( 0 < C < 1 \). The error is multiplied by \( C \) each iteration.
- **Quadratic**: \( p = 2 \). Roughly a doubling of the correct digits each iteration.
- **Superlinear**: \( p = 1, C = 0 \). “Faster than linear”. Includes quadratic convergence but also “intermediate” rates.
Taylor series with remainder term

Let \( f : \mathbb{R}^n \to \mathbb{R} \) be of class \( C^2 \) (twice continuously differentiable). Then, \( \forall x, y \in \mathbb{R}^n \),

\[
f(y) = f(x) + (y - x)^T \nabla f(x) + \frac{1}{2} (y - x)^T \nabla^2 f(\xi) (y - x)
\]

where \( \xi = \alpha y + (1 - \alpha) x \) for some \( \alpha \in [0, 1] \).

The \textbf{gradient} of \( f \):

\[
\nabla f = \left( \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)^T
\]

The \textbf{Hessian} of \( f \):

\[
\nabla^2 f = 
\begin{pmatrix}
\frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2}
\end{pmatrix}
\]
Positive definite matrices

An \( n \)-by-\( n \) real matrix \( A \) is positive semidefinite if

\[
v^T A v \geq 0 \quad \forall v \in \mathbb{R}^n
\]

It is positive definite if

\[
v^T A v > 0 \quad \forall 0 \neq v \in \mathbb{R}^n
\]

- A positive definite matrix is nonsingular
- Matrix \( A \) is positive definite if and only if matrix \( A^{-1} \) is positive definite

A symmetric matrix \( A \) is positive definite if and only if

- All eigenvalues of \( A \) are strictly positive
- \( A = LL^T \) with \( L \) lower triangular and \( L_{ii} > 0 \) (Cholesky factorization).
Unconstrained minimization: necessary and sufficient conditions

**First-order necessary condition:** Assume $f : \mathbb{R}^n \to \mathbb{R}$ has a local minimum at $x = x^*$ and that $f$ is differentiable at $x = x^*$. Then $\nabla f(x^*) = 0$.

**Second-order necessary condition:** Assume $f : \mathbb{R}^n \to \mathbb{R}$ has a local minimum at $x = x^*$ and that $f$ is of class $C^2$. Then $\nabla^2 f(x^*)$ is positive semidefinite.

**Second-order sufficiency condition:** Assume $f : \mathbb{R}^n \to \mathbb{R}$ is of class $C^2$, $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite then $x^*$ is a local minimizer.