Week 4, Lecture 1

Tuesday 6 February 2024 15:59

In this lecture, we look at Chapter 6 of the typed notes, which deals with **Convergence Analysis** of Line-search methods. The starting-point is a generic Line-search method:

$$\underline{x}_{k+1} = \underline{x}_k + \alpha_k p_k .$$

We would like:

$$\underline{x}_k \longrightarrow \underline{x}_* \quad \text{as} \quad k \longrightarrow \infty . \qquad (1)$$

To investigate the circumstances in which such **convergence** is achieved, we need some notation:

$$\boxed{\cos \theta_k = -\frac{\langle p_k, \nabla f_k \rangle}{\|p_k\|_2 \, \|\nabla f_k\|_2}} .$$

Also,

$$\nabla f_k \equiv \nabla f(\underline{x}_k)$$

In this lecture, we don't actually prove the convergence result (1) (that will come later). Instead, we look at a key intermediate result, which will help us to establish (1).

Theorem (Zoutendijk's Condition):

Consider an iterative LS method $\underline{x}_{k+1} = \underline{x}_k + \alpha_k \underline{p}_k$ where $\underline{p}_k$ is a descent direction, $\nabla f_k \cdot p_k < 0$.

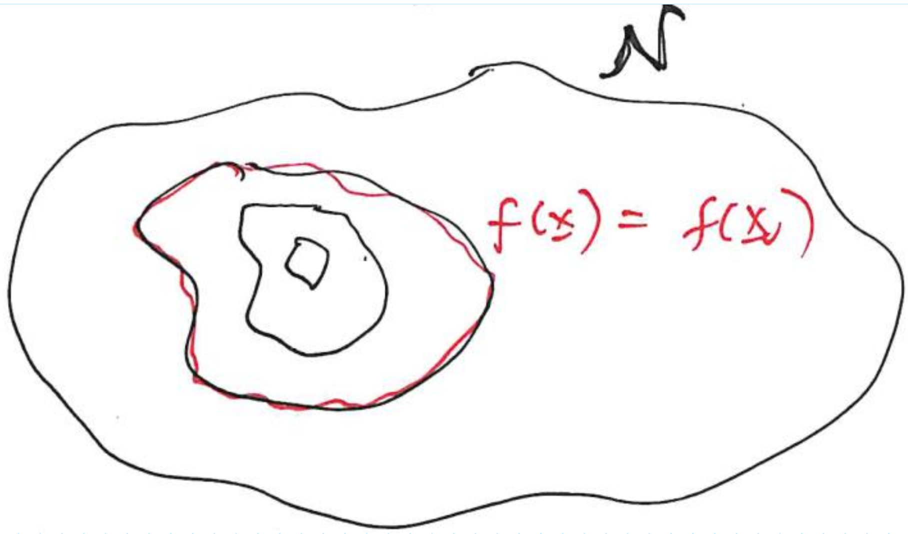
Suppose that α_k satisfies the SWCs.

Suppose also:

1. f is bounded below in $\mathbb{R}^n$
2. f is continuously differentiable in an open set $\mathcal{N}$ containing the level sets
$$\mathcal{L} = \{ \underline{x} \mid f(\underline{x}) \le f(\underline{x}_0) \}$$
where $\underline{x}_0$ is the starting-value in the iterative method
3. ∇f is Lipschitz in $\mathcal{N}$, i.e. there exists a constant $L > 0$ such that:
$$\|\nabla f(\underline{x}) - \nabla f(\underline{y})\|_2 \le L \|\underline{x} - \underline{y}\|_2$$
for all $\underline{x}, \underline{y}$ in $\mathcal{N}$.

Then:
$$\sum_{k \geq 0} \cos^2\theta_k \, \|\nabla f_k\|_2^2 < \infty. \qquad \text{(2)}$$

For the idea behind Condition 2, see the figure:



$f(\underline{x}_0) = \text{Const}.$

The proof of Zoutendijk's Theorem is in the lecture notes: we won't go into it here. Instead, we will focus our efforts on proving the following Corollary:

<u>Corollary 6.1</u> If the conditions in Zoutendijk's theorem are satisfied, then

$$\cos^2\theta_k \, \|\nabla f_k\|_2^2 \longrightarrow 0$$

as $k \longrightarrow \infty$.

<u>Proof</u>: By the assumptions in Corollary 6.1, Equation (2) is true. Hence

$$\sum_{k=0}^{\infty} \cos^2\theta_k \|\nabla f_k\|_2^2$$

is a convergent series. Hence, the general term in the sequence

$$\cos^2\theta_0 \, \|\nabla f_0\|_2^2 ,$$
$$\cos^2\theta_1 \, \|\nabla f_1\|_2^2 , \ldots$$
$$\cos^2\theta_k \, \|\nabla f_k\|_2^2 , \ldots$$

tends to zero, as $k \to \infty$:

$$\cos^2\theta_k \, \|\nabla f_k\|_2^2 \longrightarrow 0 \text{ as } k \to \infty.$$

Corollary 6.1 now has the following import consequence:

If we can "keep $\cos\theta_k$ away from zero in the tail of the sequence", i.e. if there exists a $\delta > 0$ and a $k_0 \in \mathbb{N}$ such that

$$|\cos\theta_k| > \delta \quad \forall \quad k \geq k_0$$

then $\|\nabla f_k\|_2^2 \longrightarrow 0$ as $k \longrightarrow \infty$.

Then, the iterative method converges (First-order optimality satisfied as $k \to \infty$).

Example: SD: $\cos\theta_k = +1$, for all k.
So once f satisfies the criteria in Zoutendijk's theorem, the SD method is guaranteed to converge.

Zoutendijk's Theorem can also be applied to Quasi-Newton methods, where the descent direction is defined by:

$$B_k \boldsymbol{p}_k = -\nabla f_k, \qquad (3)$$

provided the matrix B_k satisfies certain sensible conditions. This is made clear in the following Theorem:

Theorem 6.2

Theorem: Consider an iterative method where the descent direction is given by (3). Suppose that B_k satisfies:

- B_k symmetric positive-definite
- $\|B_k\|_2 \, \|B_k^{-1}\|_2 \le M$, $M = \text{const.}$, true for all k.

Then $\cos\theta_k \ge 1/M$, for all k.

We will look at the proof of this statement in the next lecture.