NOTES ON CHAPTER 3

1. Set-up

In what follows, $f : \mathbb{R}^2 \to \mathbb{R}$ is continuous function, $a = (x_0, y_0) \in \mathbb{R}^2$ is a fixed point in $\mathbb{R}^2$.

2. First Taylor formula

If f is differentiable at a, there exists a function $\omega_1(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{R}$ (depending on a and f) such that

$$\begin{cases} f(\mathbf{x}) = f(a) + \nabla f(a) \cdot (\mathbf{x} - a) + \|\mathbf{x} - a\| \, \omega_1(\mathbf{x}) \\ \lim_{\mathbf{x}\to a} \omega_1(\mathbf{x}) = 0 \end{cases}$$

This is an immediate consequence of the fact that f is differentiable at a, which is the statement that $\lim_{\mathbf{x}\to a} \frac{f(\mathbf{x}) - f(a) - \nabla f(a)\cdot(\mathbf{x}-a)}{\|\mathbf{x}-a\|} = 0$.

2.1. **Corollary: first order approximation.** For $\mathbf{x} = (x, y)$ near $a = (x_0, y_0)$,

$$f(\mathbf{x}) \approx f(a) + \nabla f(a) \cdot (\mathbf{x} - a)$$

in other words

$$f(x, y) \approx f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

3. Second Taylor formula

If f is a $\mathbf{C^2}$ function (it has second order partial derivatives and these are continuous) then there exists a (second order error) function $\omega_2(\mathbf{x})$ (depending on f and a) such that

$$\begin{cases} f(\mathbf{x}) = f(a) + \nabla f(a) \cdot (\mathbf{x} - a) + \frac{1}{2} H_f(a)[\mathbf{x} - a] + \|\mathbf{x} - a\|^2 \, \omega_2(\mathbf{x}) \\ \lim_{\mathbf{x}\to a} \omega_2(\mathbf{x}) = 0 \end{cases}$$

where:

$$H_f(a) = \begin{bmatrix} f_{xx}(a) & f_{xy}(a) \\ f_{yx}(a) & f_{yy}(a) \end{bmatrix}$$

is the *Hessian of f at a*, and notice that it is a symmetric matrix. Also, the square-bracket action is given by

$$H_f(a)[\mathbf{x} - a] = H_f(a)(\mathbf{x} - a) \cdot (\mathbf{x} - a)$$

Note: The proof of the second Taylor formula is not as straight-forward as the one for the first Taylor formula. Consult the textbook for details.

3.1. **Corollary: second order approximation.** Let $\mathbf{x} = (x, y)$ near $a = (x_0, y_0)$. Unravelling the term involving the Hessian

$$\begin{aligned} H_f(a)[\mathbf{x} - a] &= [x - x_0, y - y_0] \cdot \begin{bmatrix} f_{xx}(a) & f_{xy}(a) \\ f_{yx}(a) & f_{yy}(a) \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= f_{xx}(a)(x - x_0)^2 + 2f_{xy}(a)(x - x_0)(y - y_0) + f_{yy}(a)(y - y_0)^2 \end{aligned}$$

we obtain the second order approximation near $a = (x_0, y_0)$:

$$\begin{aligned} f(x, y) &\approx f(a) + \nabla f(a) \cdot (\mathbf{x} - a) + \frac{1}{2} H_f(a)[\mathbf{x} - a] \\ &= f(a) + f_x(a)(x - x_0) + f_y(a)(y - y_0) \\ &+ \frac{1}{2} \left\{ f_{xx}(a)(x - x_0)^2 + 2f_{xy}(a)(x - x_0)(y - y_0) + f_{yy}(a)(y - y_0)^2 \right\} \end{aligned}$$

4. Extrema of real-valued functions

4.1. **Critical points. Definition** a is a critical point for the function $f : \mathbb{R}^2 \to \mathbb{R}$ if $\nabla f(a) = 0$ ($f'(a) = 0$).

4.2. **Local extrema.** $a \in \mathbb{R}^2$ is a local maximum of f if there exists a positive number $r > 0$ such that:

$$\text{for } \mathbf{x} \in B_r(a), f(\mathbf{x}) \leq f(a)$$

Similarly for local minimum.

Theorem. If a is a local extremum for f, then a is a critical point.

Consequence: when searching for local extrema we need to restrict our search to critical points.

4.3. **Classification of critical points.** Assume a is critical , $\nabla f(a) = 0$.

(1) If $\det H_f(a) > 0$ and $f_{xx}(a) > 0$ then a is a local minimum for f

(2) If $\det H_f(a) > 0$ and $f_{xx}(a) < 0$ then a is a local maximum for f

(3) If $\det H_f(a) < 0$ then a is a saddle point for f

(4) If $\det H_f(a) = 0, a$ is a *degenerate* critical point (inconclusive situation)