

**Problem 3.2.** Assume  $f \in C^2(W; \mathbb{R})$  for some open subset  $W$  of  $\mathbb{R}^n$ .

- (a) Explicitly write out all the terms of the second order Taylor polynomial  $P_2(z)$  if  $n = 1$ ,  $n = 2$ , or  $n = 3$ , by evaluating all the multi-index powers and factorials. (This is quick for  $n = 1$  and  $n = 2$  but takes a bit of writing for  $n = 3$ .)
- (b) For arbitrary  $n$ , write down a formula for  $P_2(z)$  in terms of the Jacobian and Hessian matrices.
- (c) (The Second Derivative Test.) A symmetric matrix  $A \in \mathbb{R}^{n \times n}$  is said to be *positive-definite* if  $x^T A x > 0$  for all  $x \in \mathbb{R}^n$ ,  $x \neq 0$ . Prove that if  $Jf(x_0) = 0$  and  $Hf(x_0)$  is a positive definite, then  $f$  has a local minimum at  $x_0$ . (Hint: The function  $z \mapsto H(x_0)z \cdot z$  is a continuous function from  $\mathbb{R}^n$  to  $\mathbb{R}$ ; its restriction to the (compact!) set  $S^1 = \{x \in \mathbb{R}^n : |x| = 1\}$  is of course continuous.) Similarly, show that if  $Jf(x_0) = 0$  and  $Hf(x_0)$  is *negative-definite* (you can guess the definition), then  $f$  has a local maximum at  $x_0$ .
- (d) Show that a 2 symmetric matrix  $A = (a_{i,j})_{i,j=1}^2$  is positive definite if and only if  $\det A$  and  $a_{1,1}$  are positive. Use this information to write down another version of the Second Derivative Test for a function of 2 variables.

*Proof.* We proceed with each part separately.

- (a) Recall the formulas for  $P_2(z)$ . Suppose  $n = 1$ . If  $|\alpha| = 0$ , then we have  $\alpha = (0)$ . If  $|\alpha| = 1$ , then we have  $\alpha = (1)$ . If  $|\alpha| = 2$ , then we have  $\alpha = (2)$ . Thus,

$$\begin{aligned} P_2(z) &= \sum_{k=0}^2 \sum_{|\alpha|=k} \frac{\partial^\alpha f(a)}{\alpha!} z^\alpha \\ &= f(a) + \frac{df(a)}{dz} z + \frac{1}{2} \frac{d^2 f(a)}{dz^2} z^2. \end{aligned}$$

Suppose  $n = 2$ . If  $|\alpha| = 0$ , then we have  $\alpha = (0, 0)$ . If  $|\alpha| = 1$ , then we have  $\alpha \in \{(1, 0), (0, 1)\}$ . If  $|\alpha| = 2$ , then we have  $\alpha \in \{(2, 0), (1, 1), (0, 2)\}$ . Thus,

$$\begin{aligned} P_2(z) &= \sum_{k=0}^2 \sum_{|\alpha|=k} \frac{\partial^\alpha f(a)}{\alpha!} z^\alpha \\ &= f(a) + \partial_1 f(a) z + \partial_2 f(a) z + \frac{1}{2} \partial_1^2 f(a) z^2 + \frac{1}{2} \partial_2^2 f(a) z^2 + \partial_1 \partial_2 f(a) z^2. \end{aligned}$$

Suppose  $n = 3$ . If  $|\alpha| = 0$ , then we have  $\alpha = (0, 0, 0)$ . If  $|\alpha| = 1$ , then we have  $\alpha \in \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ . If  $|\alpha| = 2$ , then we have

$$\alpha \in \{(2, 0, 0), (0, 2, 0), (0, 0, 2), (1, 1, 0), (1, 0, 1), (0, 1, 1)\}.$$

Thus,

$$\begin{aligned} P_2(z) &= f(a) + \partial_1 f(a)z + \partial_2 f(a)z + \partial_1 f(a)z \\ &\quad + \frac{1}{2}\partial_1^2 f(a)z^2 + \frac{1}{2}\partial_2^2 f(a)z^2 + \frac{1}{2}\partial_3^2 f(a)z^2 \\ &\quad + \partial_1 \partial_2 f(a)z^2 + \partial_2 \partial_3 f(a)z^2 + \partial_1 \partial_3 f(a)z^2. \end{aligned}$$

(b) Reading off the terms, we can see that

$$P_2(z) = f(a) + \nabla f(a)^T z + \frac{1}{2} z^T H f(a) z.$$

(Note how the  $1/2$  factor perfectly handles the duplicated partials in the off diagonals.)

(c) **TODO**

(d) We want to show that  $z^T A z > 0$  for all  $z \in \mathbb{R}^2$ . Let  $a_{1,1} = a$ ,  $a_{1,2} = a_{2,1} = b$ , and  $a_{2,2} = c$  and let  $z = (x, y)$ . Expanding this out, we have

$$z^T A z = ax^2 + 2bxy + cy^2.$$

( $\Rightarrow$ ): If  $z^T A z > 0$ , then setting  $z = (1, 0)$  implies  $a > 0$ . Next consider  $z = (x, 1)$ . Then we know that  $ax^2 + 2bx + c > 0$  for all  $x$ . This means that the polynomial has no zeros. Thus its discriminant  $4b^2 - 4ac$  must be negative, i.e.  $ac - b^2 = \det A > 0$ .

( $\Leftarrow$ ): Do casework on  $y = 0$  and do almost the opposite argument. For all  $z = (x, 0)$ ,  $a_{1,1} > 0$  implies  $ax^2 = z^T A z > 0$ .

If  $y \neq 0$ , then consider  $t = x/y$  and  $q(t) = at^2 + 2bt + c = \frac{1}{y^2} z^T A z$ . Since  $\det A > 0$ , we know that  $b^2 - ac < 0$ , so  $q(t)$  has no roots. Since  $a > 0$ , it must be that  $q(t) > 0$  for all  $t$ . This implies  $z^T A z > 0$  for all  $(x, y)$ ,  $y \neq 0$ .

Thus  $z^T A z$  holds for all  $z \in \mathbb{R}^2$ , as desired.

□