ARE211, Fall 2009

CALCULUS4: THU, OCT 29, 2009 PRINTED: AUGUST 27, 2009 (LEC# 19)

Contents

4. Univariate and Multivariate Differentiation (cont) 1
4.6. Taylor's Theorem (cont) 2
4.7. Applying Taylor theory: 2nd order conditions for an unconstrained local max. 2
4.8. The Hessian, the tangent plane and the graph of f 6
4.9. Sufficient conditions for quasi-concavity 8
4.10. Terminology Review 9

4. Univariate and Multivariate Differentiation (cont)

Key points:

(1) If $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable and Hf is negative definite at $\mathbf{x}$, then f attains a strict local maximum at $\mathbf{x}$ iff $\bigtriangledown f(\mathbf{x}) = 0$ (see p. 3).
(2) If we replace "negative definite at $\mathbf{x}$" with "globally negative definite" then we can replace "local maximum" with "global maximum" in
(3) For $f : \mathbb{R}^n \to \mathbb{R}$, if Hf is twice continuously differentiable and negative definite at $\mathbf{x}$, then for every $\mathbf{x} \in \mathbb{R}^n$, the tangent plane to f at $\mathbf{x}$ is locally above the graph of f. (see p. 9).
(4) If we replace "negative definite at $\mathbf{x}$" with "globally negative definite" then we can replace "is locally" with "is globally" in 3

(5) A sufficient condition for f to be quasi-concave is that at each $\mathbf{x}$ the Hessian of f is globally negative definite on the subspace that is orthogonal to $\bigtriangledown f(\mathbf{x})$.

4.6. **Taylor's Theorem (cont)**

Taylor's Theorem (continued): Why is the theorem so tremendously important? Because if you are only interested in the *sign* of $(f(\bar{\mathbf{x}}+\mathbf{dx})-f(\bar{\mathbf{x}}))$ and you have an n'th order Taylor expansion, then you know that for some neighborhood about $\bar{\mathbf{x}}$, the sign of your expansion will be the same as the sign of the true difference, i.e., $(f(\bar{\mathbf{x}}+\mathbf{dx})-f(\bar{\mathbf{x}}))$.

4.7. **Applying Taylor theory: 2nd order conditions for an unconstrained local max.**

Definition: A function $f : X \rightarrow \mathbb{R}$ attains a strict (weak) local maximum at $\mathbf{x} \in X$ if there exists $\epsilon > 0$ such that for all $\mathbf{y} \in X \cap B(\mathbf{x}, \epsilon)$, $f(\mathbf{y}) - f(\bar{\mathbf{x}}) < (\leq) \ 0$.

Going to be talking about necessary and sufficient conditions for a local optimum of a differentiable function.

Terminology is that first order conditions are *necessary* while second order conditions are *sufficient.*

The terms necessary and sufficient conditions have a formal meaning:

- If an event A cannot happen unless an event B happens, then B is said to be a *necessary condition* for A.
- If an event B *implies* that an event A will happen, then B is said to be a *sufficient condition* for A.

For example, consider a differentiable function from $\mathbb{R}^1$ to $\mathbb{R}^1$.

- f cannot attain an interior maximum at $\bar{x}$ *unless* $f'(\bar{x}) = 0$.
  - i.e., the maxmimum is A; the derivative condition is B.
  - Thus, the condition that the first derivative is zero is *necessary* for an interior maximum; called the first order conditions.
  - Emphasize strongly that this necessity business is delicate: derivative condition is only necessary provided that f is differentiable *and* we're talking interior maximum Also, only talking LOCAL maximum.
- $f'(\bar{x}) = 0$ certainly doesn't IMPLY that f attains an interior maximum at $\bar{x}$
- If $f''(\bar{x}) < 0$, then the condition $f'(\bar{x}) = 0$ is both necessary and sufficient for an interior *local* maximum;
- Alternatively, if you know in advance that f is *strictly concave*, then the condition that $f'(\bar{x})$ is zero is necessary and sufficient for a *strict global* maximum.

Generalizing to functions defined on $\mathbb{R}^n$, a simple application of Taylor's theorem proves that if an appropriate local second order condition is satisfied, then the first order conditions are in fact necessary and sufficient for a *strict local* maximum

Theorem: Let X be an open subset of $\mathbb{R}^n$ and consider a twice continuously differentiable function $f : X \to \mathbb{R}$ and a point $\bar{\mathbf{x}} \in \mathbb{R}^n$ such that the Hessian of f is negative (semi) definite at $\bar{\mathbf{x}}$. f attains a (weak) strict local maximum at $\bar{\mathbf{x}} \in \mathbb{R}^n$ iff $\bigtriangledown f(\bar{\mathbf{x}}) = 0$.

Proof of Necessity: To prove this part, we use the *local* version of Taylor's theorem. Suppose $\bigtriangledown f(\bar{\mathbf{x}}) \neq 0$. We'll show that in this case, $f(\cdot)$ cannot be maximized at $\bar{\mathbf{x}}$. Consider the first order Taylor expansion of f about $\bar{\mathbf{x}}$:

$$f(\bar{\mathbf{x}} + \mathbf{dx}) - f(\bar{\mathbf{x}}) = \bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} + \text{ a remainder term } . \tag{1}$$

Since X is open, there exists $\epsilon_1 > 0$ such that $B(\mathbf{x}, \epsilon_1) \subset X$. From the Taylor Young (Local Taylor) theorem there exists $\epsilon_2 \in (0, \epsilon_1)$ such that if $||\mathbf{dx}|| < \epsilon_2$, then the absolute value of the first term of the Taylor expansion is larger than the absolute value of the remainder term. Now, let $\mathbf{dx} = \lambda \bigtriangledown f(\bar{\mathbf{x}})$, where $\lambda > 0$ is picked sufficiently small that $||\lambda \bigtriangledown f(\bar{\mathbf{x}})|| < \epsilon_2$. Since the angle θ between $\bigtriangledown f(\bar{\mathbf{x}})$ and $\mathbf{dx}$ is zero, we have $\bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} = \cos(\theta)|| \bigtriangledown f(\bar{\mathbf{x}})||||\mathbf{dx}|| > 0$. Therefore, $f(\bar{\mathbf{x}} + \mathbf{dx}) > f(\bar{\mathbf{x}})$. Since $\bar{\mathbf{x}} + \mathbf{dx} \in B(\bar{\mathbf{x}}, \epsilon_2) \subset B(\bar{\mathbf{x}}, \epsilon_1) \subset X$, $\bar{\mathbf{x}}$ does not maximize f on X. □

Proof of joint sufficiency: To prove sufficiency, we use the *global* version of Taylor's theorem. We need to show that there exists $\epsilon > 0$ such that for all $\mathbf{dx} \in B(0, \epsilon)$, $f(\bar{\mathbf{x}} + \mathbf{dx}) < f(\bar{\mathbf{x}})$. Let S denote the unit sphere, i.e., $S = \{\mathbf{v} \in \mathbb{R}^n : ||\mathbf{v}|| = 1\}$. Define the function $\delta(\cdot)$ on S by for $\mathbf{v} \in S$, $\delta(\mathbf{v}) = \mathbf{v}'\text{Hf}(\bar{\mathbf{x}})\mathbf{v}$. Since $\text{Hf}(\bar{\mathbf{x}})$ is negative definite, $\delta(\cdot) < 0$. Since $\delta(\cdot)$ is continuous and S is compact, $\delta(\cdot)$ attains a maximum on S. Let $\underline{\delta} = 0.5 \max(\delta(\cdot)) < 0$. Since f is twice continuously differentiable, there exists $\epsilon > 0$, such that for all $\mathbf{x}' \in B(\bar{\mathbf{x}}, \epsilon)$ and all $\mathbf{v} \in S$, $\mathbf{v}'\text{Hf}(\mathbf{x}')\mathbf{v} < \underline{\delta} < 0$. Hence for an *arbitrary vector* $0 \neq \mathbf{dx} \in \mathbb{R}^n$, $\mathbf{dx}'\text{Hf}(\mathbf{x}')\mathbf{dx} < 0$. Now pick $\mathbf{dx} \in B(0, \epsilon)$. Clearly for all $\lambda \in [0, 1]$, $\lambda\mathbf{dx} \in B(0, \epsilon)$ also, so that $\text{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})$ is negative definite; in particular $\frac{1}{2}\mathbf{dx}'\text{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})\mathbf{dx} < 0$. It now follows from the Taylor Lagrange (global Taylor) theorem that for some $\lambda \in [0, 1]$,

$$\begin{aligned} f(\bar{\mathbf{x}} + \mathbf{dx}) - f(\bar{\mathbf{x}}) &= \bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} + \frac{1}{2}\mathbf{dx}'\text{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})\mathbf{dx} \qquad (2) \\ &< \bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} = 0. \end{aligned}$$

This completes the proof that $f(\bar{\mathbf{x}} + \mathbf{dx}) < f(\bar{\mathbf{x}})$. □.

One has to be extremely careful about the wording of these necessary and sufficient conditions: The following statement is FALSE: A function $f : \mathbb{R}^n \to \mathbb{R}$ attains a strict local maximum at $\bar{\mathbf{x}} \in \mathbb{R}^n$ iff $\bigtriangledown f(\bar{\mathbf{x}}) = 0$ and Hf is negative definite at $\bar{\mathbf{x}}$. The "if" part of this statement is true, but the "only if" isn't: you could have a strict local max at $\bar{\mathbf{x}}$ without f being negative definite at $\bar{\mathbf{x}}$, e.g. $-x^4$ attains a global max at 0 but it isn't negative definite at 0.

The theorem above only gives sufficient conditions for a *local* maximum. On the other hand, if $\text{Hf}(\cdot)$ is *globally* negative definite, we obtain the following global result.

Theorem: Consider a twice continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$. If $\mathrm{Hf}(\cdot)$ is globally negative (semi-)definite, f attains a strict (weak) *global* maximum at $\bar{\mathbf{x}} \in \mathbb{R}^n$ iff $\bigtriangledown f(\bar{\mathbf{x}}) = 0$.

The difference between this and the previous result (ignoring the 'semi' in parenthesees) is that that we have replaced negative definiteness at a point with global negative definiteness. This allows us to consider large $\mathbf{dx}$'s instead of just small ones, and obtain conditions for a *global* rather than just a local maximum.

Proof: If $\bigtriangledown f(\bar{\mathbf{x}}) \neq 0$, the theorem above establishes that $\mathbf{x}$ cannot be a local maximum, and hence it certainly isn't a global maximum. So necessity follows from the preceding theorem. Now suppose that $\bigtriangledown f(\bar{\mathbf{x}}) = 0$. Rewriting expression (4) above, we have that for *any* $\mathbf{dx} \in \mathbb{R}^n$

$$f(\bar{\mathbf{x}} + \mathbf{dx}) - f(\bar{\mathbf{x}}) \quad = \quad \bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} + \frac{1}{2}\mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})\mathbf{dx}, \text{ for some } \lambda \in [0,1] \tag{4}$$

which, since by assumption $\bigtriangledown f(\bar{\mathbf{x}}) = 0$,

$$= \quad \frac{1}{2}\mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})\mathbf{dx}$$

Since $\mathrm{Hf}(\cdot)$ is globally negative (semi)-definite, it follows that $f(\bar{\mathbf{x}} + \mathbf{dx}) - f(\bar{\mathbf{x}}) < (\leq) \; 0$, proving that $f(\cdot)$ is globally strictly (weakly) maximized at $\mathbf{x}$. □

A slight modification of necessity argument we used above can be used to obtain another useful result, that we'll use later to show that a definiteness properties implies quasi-concavity or quasi-convexity. Suppose we are maximizing or minimizing a function f on an *open* line segment L. A necessary condition for a point to be an extremum (i.e., maximum or minimum) on L is that the gradient of f at that point is orthogonal to the line segment. Formally:

Lemma: Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable. Fix $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and let $L = \{\lambda\mathbf{x} + (1-\lambda)\mathbf{y} : \lambda \in (0,1)\}$. A necessary condition for $\bar{\mathbf{x}} \in L$ to be a local extremum on L is that $\bigtriangledown f(\bar{\mathbf{x}}) \cdot \mathbf{dx} = 0$, for any $\mathbf{dx}$ such that $\bar{\mathbf{x}} + \mathbf{dx} \in L$.

Proof: Once again, we use the local version of Taylor's theorem. We'll deal only with local *maxima*; the argument for minima is parallel. Suppose $\bigtriangledown f(\bar{\mathbf{x}}) \cdot \mathbf{v} \neq 0$, for some $\mathbf{v}$ such that $\bar{\mathbf{x}} + \mathbf{v} \in L$. We'll show that in this case, $f(\cdot)$ cannot be locally maximized at $\bar{\mathbf{x}}$. Since L is open, there exists $\epsilon_1 > 0$

such that $||\lambda \mathbf{v}|| < \epsilon_1$ implies that $\mathbf{x} + \lambda \mathbf{v} \in L$. (Note that λ could be either positive or negative.) Now consider the first order Taylor expansion of f about $\bar{\mathbf{x}}$:

$$f(\bar{\mathbf{x}} + \mathbf{dx}) - f(\bar{\mathbf{x}}) = \bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} + \text{ a remainder term }. \tag{3}$$

From the Taylor Young (Local Taylor) theorem there exists $\epsilon_2 \in (0, \epsilon_1)$ such that if $||\mathbf{dx}|| < \epsilon_2$, then the absolute value of the first term of the Taylor expansion is larger than the absolute value of the remainder term. Now pick $\lambda \in \mathbb{R}$, such that $||\lambda \mathbf{v}|| < \epsilon_2$ and $\bigtriangledown f(\bar{\mathbf{x}}) \cdot \lambda \mathbf{v} > 0$. (If $\bigtriangledown f(\bar{\mathbf{x}}) \cdot \mathbf{v} < 0$, obviously λ will have to be negative.) Since $||\mathbf{dx}|| < \epsilon_2$, $f(\bar{\mathbf{x}} + \mathbf{dx}) > f(\bar{\mathbf{x}})$. Since $\bar{\mathbf{x}} + \mathbf{dx} \in L$, we have established that $\bar{\mathbf{x}}$ does not locally maximize f on L. □

4.8. **The Hessian, the tangent plane and the graph of f**

Another important application of Taylor's theorem is the following result, which is obtained by essentially duplicating the proof of sufficiency above.

Theorem: Given $f : \mathbb{R}^n \to \mathbb{R}$ twice continuously differentiable and $\bar{\mathbf{x}} \in \mathbb{R}^n$, if Hf is positive (negative) definite at $\bar{\mathbf{x}}$ then there exists $\epsilon > 0$ such that the tangent plane to f at $\bar{\mathbf{x}}$ lies *below* (*above*) the graph of the function on the ϵ-ball around $\bar{\mathbf{x}}$.

Proof: We'll do the case when Hf is negative definite at $\bar{\mathbf{x}}$. Once again, for any $\mathbf{dx} \neq 0$,

$$f(\bar{\mathbf{x}} + \mathbf{dx}) - f(\bar{\mathbf{x}}) = \bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx} + \frac{1}{2}\mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})\mathbf{dx}, \text{ for some } \lambda \in [0,1] \tag{4}$$

Now, since $\mathrm{Hf}(\bar{\mathbf{x}})$ is negative definite by assumption, and $\mathrm{Hf}(\cdot)$ is continuous, there exists $\epsilon > 0$ such that if $\mathbf{dx} \in B(0, \epsilon)$ and $\lambda \in [0,1]$, then $\mathrm{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})$ will be negative definite also. Hence the second term in the expression above will be negative. Subtracting $\bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx}$ from both sides, we obtain that for all $\mathbf{dx}$ with $||\mathbf{dx}|| < \underline{\epsilon}$,

$$\underbrace{f(\bar{\mathbf{x}} + \mathbf{dx})}_{\text{the height of } f \text{ at } \bar{\mathbf{x}} + \mathbf{dx}} - \underbrace{\left(f(\bar{\mathbf{x}}) + \bigtriangledown f(\bar{\mathbf{x}})\mathbf{dx}\right)}_{\text{the height of the tangent plane at } \bar{\mathbf{x}} + \mathbf{dx}} = \frac{1}{2}\mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})\mathbf{dx} < 0 \tag{5}$$

□

Notice that the theorem above isn't necessarily true without restricting $\mathbf{dx}$ to lie in a neighborhood of zero. Think of a camel. Put a tangent plane against the smaller hump, and the whole camel isn't underneath the plane. On the other hand, if Hf(·) is *globally* negative definite, we can prove a stronger result.

Theorem: If $f : \mathbb{R}^n \to \mathbb{R}$ if twice continuously differentiable and Hf(·) is negative (semi) definite on its entire domain, then for all $\mathbf{x} \in \mathbb{R}^n$, the tangent plane to f at $\mathbf{x}$ lies everywhere (weakly above) above the graph of the function.

Proof: The proof is identical to the proof of the preceding result except that we omit the caveat about $\mathbf{dx}$ being small. Since Hf(·) is *everywhere* negative (semi) definite, then we know that regardless of the size of $\mathbf{dx}$, the matrix $\mathrm{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})$ is, in particular, negative (semi) definite, so that the term $\frac{1}{2}\mathbf{dx}'\mathrm{Hf}(\bar{\mathbf{x}} + \lambda\mathbf{dx})\mathbf{dx}$ in expression (4) above is negative (non-positive). Hence the left hand side of expression (5) above is also negative (non-positive). □

Finally, we prove that for twice-differentiable functions, concavity and global negative-semi-definiteness of the Hessian are equivalent conditions. Theorem: If $f : \mathbb{R}^n \to \mathbb{R}$ if twice continuously differentiable then f is concave if and only if Hf(·) is negative semi-definite on its entire domain.

Proof: The necessity part follows easily from the above theorem. I'll leave it as an exercise. Sufficiency is harder. We'll prove that if f is not concave then $Hf(\cdot)$ is not globally semi-definite. Suppose that f is not concave, i.e., that there exists $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ and $\lambda' \in (0,1)$ such that $f(\lambda'\mathbf{y} + (1-\lambda')\mathbf{z}) < \lambda' f(\mathbf{y}) + (1-\lambda')f(\mathbf{z})$. For each $\lambda \in [0,1]$, let $g(\lambda) = f(\lambda\mathbf{y} + (1-\lambda)\mathbf{z}) - \big[\lambda f(\mathbf{y}) + (1-\lambda)f(\mathbf{z})\big]$. Note that $g(0) = g(1) = 0$, while by assumption $g(\lambda') < 0$. Moreover, since the term in square brackets is affine, it does not contribute to $g''(\lambda)$. Indeed, applying the chain rule.

$$g''(\lambda) \quad = \quad \frac{d^2}{d\lambda^2}\big(f(\lambda\mathbf{y} + (1-\lambda)\mathbf{z})\big) \quad = \quad (\mathbf{y} - \mathbf{z})' Hf(\lambda\mathbf{y} + (1-\lambda)\mathbf{z})(\mathbf{y} - \mathbf{z})$$

Since $[0,1]$ is compact and g is continuous, g attains a global minimum on $[0,1]$. Let $\bar{\lambda}$ be a minimizer of g on $[0,1]$. Since $g(\bar{\lambda}) \leq g(\lambda') < 0$, $\bar{\lambda} \in (0,1)$. Hence $g'(\bar{\lambda}) = 0$. Let $d\lambda = -\bar{\lambda}$ so that $\bar{\lambda} + d\lambda = 0$. By Global Taylor, there exists $\lambda \in [0, \bar{\lambda}]$ such that

$$g(0) - g(\bar{\lambda}) = g'(\bar{\lambda})d\lambda + 0.5g''(\lambda)d\lambda^2 = 0.5g''(\lambda)d\lambda^2 = 0.5d\lambda^2(\mathbf{y} - \mathbf{z})' Hf(\lambda\mathbf{y} + (1-\lambda)\mathbf{z})(\mathbf{y} - \mathbf{z}) > 0$$

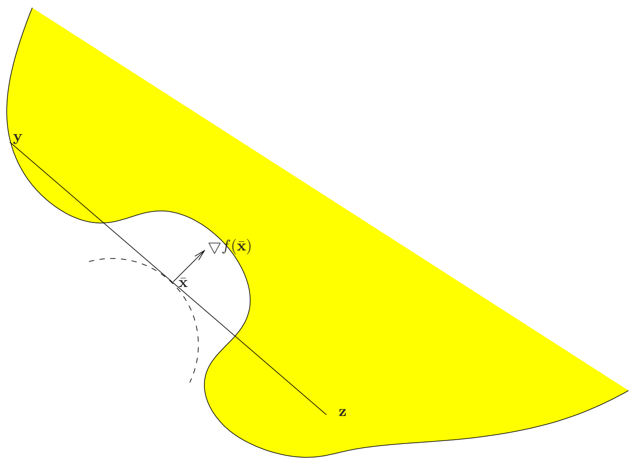

FIGURE 1. Negative definiteness subject to constraint implies QC

Hence we have established that $Hf(\cdot)$ is not negative semi-definite at $\lambda\mathbf{y} + (1-\lambda)\mathbf{z}$. □

4.9. **Sufficient conditions for quasi-concavity**

We can use an analogous argument to show that "negative definiteness subject to constraint" implies quasi-concavity. Recall that a *sufficient* condition for a twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ to be strictly quasi-concave is that for all $\mathbf{x}$, $\mathrm{Hf}(\mathbf{x})$ is negative definite *on the subspace of* $\mathbb{R}^n$ *which is orthogonal to the gradient of* f. Formally the sufficient condition is:

$$\text{for all } \mathbf{x} \text{ and all } \mathbf{dx} \neq 0 \text{ such that } \bigtriangledown f(\mathbf{x}) \cdot \mathbf{dx} = 0, \mathbf{dx}'\mathrm{Hf}(\mathbf{dx})\mathbf{dx} < 0. \tag{6}$$

Theorem: If $f : \mathbb{R}^n \to \mathbb{R}$ is thrice continuously differentiable, then (6) is a sufficient condition for f to be quasi-concave.

Recall that f is quasi-concave if all upper contour sets are convex. We'll prove that if f is *not* quasi-concave then (6) is violated. This is a little tricky to do: we have to find a point $\mathbf{x}$ and a movement $\mathbf{dx}$ such that $\bigtriangledown f(\mathbf{x}) \cdot \mathbf{dx} = 0$ but $\mathbf{dx}'\mathrm{Hf}(\mathbf{x})\mathbf{dx} \geq 0$. Fig. 1 shows how to find it: we find the *minimum* $\bar{\mathbf{x}}$ of f on the line segment joining $\mathbf{y}$ and $\mathbf{z}$, and then invoke the Lemma above, which says that the gradient of f at $\bar{\mathbf{x}}$ must be orthogonal to the line segment. Then, we can use local Taylor to establish that the level set corresponding to $f(\bar{\mathbf{x}})$ *must, locally,* look something like the dashed line, and we're done. The above isn't quite precise, but the proof below is.

Proof: Assume that f is not quasi-concave. In this case, there must exist $\mathbf{y}, \mathbf{z}, \mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x} = \lambda\mathbf{y} + (1-\lambda)\mathbf{z}$, for some $\lambda \in (0,1)$, and $f(\mathbf{x}) < f(\mathbf{y}) \leq f(\mathbf{z})$. We will now show that property (6) is not satisfied. Let $\bar{L} = \{\lambda\mathbf{y} + (1-\lambda)\mathbf{z} : \lambda \in [0,1]\}$ and let L denote the interior of $\bar{L}$. (Note that $\bar{L}$ is a *closed* interval, cf the open interval L in the Lemma above.) Since f is continuous and $\bar{L}$ is compact, f attains a (global) minimum on $\bar{L}$. Clearly, since by assumption $f(\mathbf{x}) < \min[f(\mathbf{y}), f(\mathbf{z})]$, it follows that the global *minimizer* $\bar{\mathbf{x}}$ of f on $\bar{L}$ necessarily belongs also to L. Since L is open, we can pick $\mathbf{v} \neq 0$ such that $\bar{\mathbf{x}} + \mathbf{v} \in L$. From the Lemma above, $\bigtriangledown f(\bar{\mathbf{x}})\mathbf{v} = 0$. There are now two possibilities: either (a) $\mathbf{v}'\mathrm{Hf}(\bar{\mathbf{x}})\mathbf{v}$ equals zero or (b) it does not. If case (a) holds, then property (6) is violated. If case (b) holds, then for all $m \in \mathbb{N}$, $T^2(f, \mathbf{x}, \mathbf{v}/m) = \frac{1}{2}\mathbf{v}'\mathrm{Hf}(\bar{\mathbf{x}})\mathbf{v}/m^2 \neq 0$. (Recall that $T^2(f, \mathbf{x}, \mathbf{v}) = \sum_{k=1}^{2} \mathrm{Tf}^{\,k}(\mathbf{x}, \mathbf{v})/k!$, where $\mathrm{Tf}^{\,k}(\mathbf{x}, \mathbf{v})$ was defined in lecture Calculus3. By Local Taylor

$$f(\bar{\mathbf{x}} + \mathbf{v}/m) - f(\bar{\mathbf{x}}) \quad = \quad T^2(f, \mathbf{x}, \mathbf{v}/m) + \text{a remainder term that depends on } m.$$

Moreover, if m is sufficiently large, $f(\bar{\mathbf{x}} + \mathbf{v}/m) - f(\bar{\mathbf{x}})$ has the same sign as $T^2(f, \mathbf{x}, \mathbf{v}/m)$. But since $\bar{\mathbf{x}}$ minimizes f on $\bar{L}$, $f(\bar{\mathbf{x}} + \mathbf{v}/m) - f(\bar{\mathbf{x}})$ cannot be positive, and hence must be negative, once again violating (6). We have thus shown that property (6) fails in both case (a) and case (b), establishing that (6) is sufficient for quasi-concavity. □

Negative definiteness is a stronger condition than we need to establish quasi-concavity. The following result is establishes the relationship between quasi-concavity and negative *semi*-definiteness.

Theorem: If $f : \mathbb{R}^n \to \mathbb{R}$ is thrice continuously differentiable, then f is quasi-concave iff for all $\mathbf{x}$ and all $\mathbf{dx}$ such that $\bigtriangledown f(\mathbf{x}) \cdot \mathbf{dx} = 0$, $\mathbf{dx}'\mathrm{Hf}(\mathbf{dx})\mathbf{dx} \leq 0$.

It's easy to prove the necessity part of this result, but harder to prove sufficiency. We won't try.

4.10. **Terminology Review**

Keep these straight.

- The derivative of f at $\bar{\mathbf{x}}$: this is a point (which may be a scalar, vector or matrix).

- partial derivative of f w.r.t variable i at $\bar{\mathbf{x}}$.
- directional derivative of f in the direction h at $\bar{\mathbf{x}}$.
- crosspartial derivative of f_i w.r.t variable j at $\bar{\mathbf{x}}$.
- total derivative of f w.r.t variable i at $\bar{\mathbf{x}}$. The total derivative is different from the partial derivative w.r.t. i iff other variables change as the i'th variable changes. In this case, the change in the other variables determine a *direction*; e.g., if q depends on p, then a change in p induces a change in (p, q)-space in the direction $h = (1, q'(p))$. However, whenever $q'(p) \neq 0$, the total derivative is *strictly larger* than the directional derivative in the direction h.

- The derivative of f: this is a function. It's the generic term
  - functions from $\mathbb{R}^1$ to $\mathbb{R}^1$: derivative is usually just called the derivative.
  - functions from $\mathbb{R}^n$ to $\mathbb{R}^1$: derivative is usually called the gradient.
  - functions from $\mathbb{R}^n$ to $\mathbb{R}^m$: derivative is called the Jacobian matrix
  - the Hessian of f is the Jacobian of the derivative of f: i.e., the matrix of 2nd partial derivatives.
- The differential of f at $\bar{x}$: this is also a function, but a different one from the gradient. It's the unique linear function whose coefficient vector is the gradient vector evaluated at $\bar{x}$. Again, the differential may be a linear mapping from $\mathbb{R}^1$ to $\mathbb{R}^1$, from $\mathbb{R}^n$ to $\mathbb{R}^1$, or from $\mathbb{R}^n$ to $\mathbb{R}^m$, depending on the domain and range of the original function f.