Lecture 26 - Application of the Hessian to local extrema

December 5, 2022

$Hf(\vec{x}) \to$ matrix of 2^{nd} partials

ij-entry is $f_{x_i x_j}$

Goals: Be able to classify critical points of a multivariable scalar function.

We will further elaborate on the second derivative test introduced in the previous lecture, since we only discussed it for functions $f : \mathbb{R}^2 \to \mathbb{R}$. Recall that the quadratic approximation of a function $f : \mathbb{R}^n \to \mathbb{R}$ at a critical point $\mathbf{a}$ is

$$f(\mathbf{a}+\mathbf{h}) \approx f(\mathbf{a}) + \frac{1}{2}\mathbf{h}^\top Hf(\mathbf{a})\mathbf{h} = f(\mathbf{a}) + \frac{1}{2}q_{Hf(\mathbf{a})}(\mathbf{h}),$$

since $\nabla f(\mathbf{a}) = \mathbf{0}$. For a single variable function $f : \mathbb{R} \to \mathbb{R}$, this says

$$f(a+h) \approx f(a) + f''(a)h^2.$$

So if $f''(a) > 0$, then $f(a+h) > f(a)$, and so a is a local minimum. Similarly, if $f''(a) < 0$, then $f(a+h) < f(a)$ and so a is a local maximum. If $f''(a) = 0$, then the above approximation tells us nothing about whether the critical point is a local extrema. Before we generalize this to higher dimensions, let us more properly define a saddle point.

<u>Definition:</u> For $n > 1$ and a function $f : \mathbb{R}^n \to \mathbb{R}$, a critical point $\mathbf{a} \in \mathbb{R}^n$ is a **saddle point** of f if there are different lines L and L' in $\mathbb{R}^n$ through $\mathbf{a}$ so that f evaluated on the line L has a local maximum at $\mathbf{a}$ and f evaluated on the line L' has a local minimum at $\mathbf{a}$.

IL $L = \vec{a} + t\vec{v}$, look at $f(\vec{a} + t\vec{v}) \to$ local max @ $t=0$

$f(\vec{a} + t\vec{v}') \longrightarrow$ local min @ $t=0$

<u>Theorem 26.1.5</u>: (Second Derivative Test, Version 1) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, and $\mathbf{a} \in \mathbb{R}^n$ a **critical point** of f.

$q_{Hf(\vec{a})}(\vec{x}) > 0$ for $x \neq 0$

1. If the Hessian $Hf(\mathbf{a})$ is positive-definite, then $\mathbf{a}$ is a local minimum of f.
2. If the Hessian $Hf(\mathbf{a})$ is negative-definite, then $\mathbf{a}$ is a local maximum of f.
3. If the Hessian $Hf(\mathbf{a})$ is indefinite, then $\mathbf{a}$ is a saddle point of f, hence neither a local min nor a local max.

If the Hessian $Hf(\mathbf{a})$ is positive/negative-semidefinite, then we need more information to conclude the behavior of $\mathbf{a}$.

Example 1: Suppose $f(x, y, z)$ has a critical point at $(1, 1, 0)$ and $Hf(1, 1, 0) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 11 \end{bmatrix}$. What type of critical point is $(1, 1, 0)$?

$$q_{Hf(1,1,0)} = 1x^2 + 5y^2 + 11z^2 + 0xy + 0yz + 0xz$$

$$= x^2 + 5y^2 + 11z^2 \quad > 0 \text{ away from } (0,0,0)$$

so positive-definite $\Rightarrow$ local min @ $(1,1,0)$

Example 2: Let $f(x, y) = -2x^2 - \frac{1}{2}y^2 - xy$. Show that f has a critical point at $(0, 0)$ and classify it.

$$\nabla f = \begin{bmatrix} -4x - y \\ -y - x \end{bmatrix} = \vec{0} \longrightarrow \begin{matrix} -4x = y \\ y = -x \end{matrix} \to -4x = -x \to x = 0,\ y = 0$$

$\Rightarrow (0,0)$ is crit point

$$Hf(x,y) = \begin{bmatrix} -4 & -1 \\ -1 & -1 \end{bmatrix} \quad \lambda^2 - \mathrm{tr}(A) + \det(A) \to \lambda^2 + 5\lambda + 3 = 0 \Rightarrow \lambda = \frac{-5 \pm \sqrt{13}}{2}$$

Both eigenvalues are negative

$\Rightarrow 0$ is a local max

Example 3: Let $f(x,y) = x^4 + y^4$, $g(x,y) = x^4 - y^4$, and $h(x,y) = -x^4 - y^4$. What type of critical point is $(0,0)$ for f, g, and h?

$\nabla f = \begin{pmatrix} 4x^3 \\ 4y^3 \end{pmatrix}$, $Hf(x,y) = \begin{pmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{pmatrix}$, $H_g(x,y) = \begin{pmatrix} 12x^2 & 0 \\ 0 & -12y^2 \end{pmatrix}$,

$Hh(x,y) = \begin{pmatrix} -12x^2 & 0 \\ 0 & 12y^2 \end{pmatrix}$

$Hf(0,0) = H_g(0,0) = Hh(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ $\therefore$ 2nd derivative test is inconclusive

---

$f(x,y) = x^4 + y^4 \geq 0$, equals 0 at $f(0,0)$ $\Rightarrow$ $(0,0)$ is a local minimum

$h(x,y) = -x^4 - y^4 \leq 0$, equals 0 at $h(0,0)$ $\Rightarrow$ $(0,0)$ is a local maximum

$g(x,y) = x^4 - y^4$ can be both pos & neg

Let $L = \{t\binom{1}{0}, t \in \mathbb{R}\}$ (x-axis), $g(L) \to g(t\binom{1}{0}) = g(t,0) = t^4 \geq 0$ local min on x-axis $\Rightarrow$ saddle point

Let $L' = \{t\binom{0}{1}, t \in \mathbb{R}\}$ (y-axis), $g(L') \to g(t\binom{0}{1}) = g(0,t) = -t^4 \leq 0$ local max on y-axis

Recall from lecture 24 that we can rewrite a quadratic form using the diagonalization formula, and so

$$q_{Hf(\mathbf{a})}(\mathbf{v}) = \sum_{i=1}^{n} \lambda_i (\mathbf{w}_i \cdot \mathbf{w}_i) t_i^2,$$

where $\mathbf{v} = \sum_{i=1}^{n} t_i \mathbf{w}_i$, and the $\mathbf{w}_i$ are the eigenvectors of $Hf(\mathbf{a})$ with eigenvalues λ_i respectively. This gave us a different way to think about definiteness (Proposition 24.2.10), and thus we can show a different version of the second derivative test:

__**Theorem 26.3.1**__: (Second Derivative Test, Version II) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, and $\mathbf{a} \in \mathbb{R}^n$ a **critical point** of f.

1. If the Hessian $Hf(\mathbf{a})$ has all eigenvalues positive, then $\mathbf{a}$ is a local minimum of f.
2. If the Hessian $Hf(\mathbf{a})$ has all eigenvalues negative, then $\mathbf{a}$ is a local maximum of f.
3. If the Hessian $Hf(\mathbf{a})$ has both positive and negative eigenvalues, then $\mathbf{a}$ is a saddle point of f, hence neither a local min nor a local max.

We have the following useful shortcut for function $f : \mathbb{R}^2 \to \mathbb{R}$.

__**Theorem 26.3.3**__: Consider a 2×2 matrix $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ (symmetric), so the eigenvalues are the roots of the characteristic polynomial $P_A(\lambda) = \lambda^2 - \mathrm{tr}(A)\lambda + \det(A)$. For its two roots $\lambda_1, \lambda_2 \in \mathbb{R}$:

1. λ_1, λ_2 have opposite signs when their product $\det(A) = \lambda_1\lambda_2$ is negative, so the indefinite case occurs exactly when $ad - b^2 = \det(A) < 0$;
2. λ_1, λ_2 are either both positive or both negative when their product $\det(A) = \lambda_1\lambda_2$ is positive, so A is positive-definite or negative-definite precisely when $ad - b^2 = \det(A) > 0$;
3. in the case of (ii), the common sign of λ_1 and λ_2 is the same as that of their sum $\lambda_1 + \lambda_2 = \mathrm{tr}(A) = a + d$.

Thus, we do not actually have to compute eigenvalues for 2×2 matrices to classify the critical points (though we do if we, for example, want a sketch of the contours of f near that critical point).

$\lambda^2 - \mathrm{tr}(A) + \det(A) = (\lambda - \lambda_1)(\lambda - \lambda_2)$

$= \lambda^2 - (\lambda_1 + \lambda_2) + \lambda_1\lambda_2$

$\mathrm{tr}(A) = \lambda_1 + \lambda_2$

$\det(A) = \lambda_1 \lambda_2$

Example 4: Redo Example 2 without computing eigenvalues of $Hf(0,0)$.

$f(x,y) = -2x^2 - \frac{1}{2}y^2 - xy$

$Hf(0,0) = \begin{bmatrix} -4 & -1 \\ -1 & -1 \end{bmatrix}$

$\det(Hf(0,0)) = 4 - 1 = 3 > 0 \therefore Hf(0,0)$ is either pos or neg definite

$\operatorname{tr}(Hf(0,0)) = -5 \Rightarrow \lambda_1 + \lambda_2 < 0$, sign of λ_1 is neg, λ_2 is neg

$\Rightarrow Hf(0,0)$ is negative definite $\Rightarrow (0,0)$ is local max

Example 5: Find the critical points of $f(x,y) = 3x^2y + 2y^3 - xy$ and classify them each as a local max, local min, or saddle point in two different ways (Theorem 26.3.1 and Theorem 26.3.3).

$\nabla f = \begin{bmatrix} 6xy - y \\ 3x^2 + 6y^2 - x \end{bmatrix} = \vec{0} \rightarrow$

$6xy = y \rightarrow y(6x-1) = 0 \rightarrow y = 0$ or $x = \frac{1}{6}$

$3x^2 + 6y^2 - x = 0$

$y = 0$: $3x^2 - x = 0$, $x(3x-1) = 0$, $x = 0, 1/3$

$x = 1/6$: $3\left(\frac{1}{36}\right) + 6y^2 - \frac{1}{6} = 0$, $y = \pm \frac{1}{6\sqrt{2}}$

$\therefore$ In total: 4 critical points: $(0,0)$, $(1/3, 0)$, $\left(1/6, \frac{1}{6\sqrt{2}}\right)$, $\left(1/6, -\frac{1}{6\sqrt{2}}\right)$

$Hf(x,y) = \begin{bmatrix} 6y & 6x-1 \\ 6x-1 & 12y \end{bmatrix}$

$(0,0)$: $Hf(0,0) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ $\det Hf(0,0) = -1 < 0 \Rightarrow \lambda_1, \lambda_2$ has opp signs $\Rightarrow$ saddle point

or $\lambda^2 - 0 - 1 = 0 \Rightarrow \lambda = \pm 1$

$(1/3, 0)$: $Hf(1/3, 0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\det Hf(1/3, 0) = -1 < 0 \Rightarrow$ saddle point

or $\lambda^2 - 0 - 1 = 0 \rightarrow \lambda = \pm 1$

$(1/6, 1/6\sqrt{2})$: $Hf(1/6, 1/6\sqrt{2}) = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$ $\lambda_1 = \sqrt{2}$, $\lambda_2 = \frac{1}{\sqrt{2}}$ $\Rightarrow$ both pos $\Rightarrow$ local min

$(1/6, -1/6\sqrt{2})$: $Hf(1/6, -1/6\sqrt{2}) = \begin{bmatrix} -1/\sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix}$ $\lambda_1 = -\sqrt{3}$ $\lambda_2 = -\frac{1}{\sqrt{2}}$ $\Rightarrow$ both neg $\Rightarrow$ local max

If there's time, here is an example that summarizes the main ideas from Chapters 25 and 26.

Example 6: For $f(x,y) = 2x^3 + xy^2 + 5x^2 + y^2$:

a • Find the quadratic approximation to f at the point $(1,0)$. $\to f(a+h) \approx f(a) + \nabla f(a)\cdot h + \frac{1}{2} h^T Hf(a)\, h$

b • Find all critical points of f.

c • Classify them using the Hessian matrix.

d • Sketch the approximate contours of f near each critical point.

a) $\nabla f(x,y) = \begin{bmatrix} 6x^2 + y^2 + 10x \\ 2yx + 2y \end{bmatrix}$, $Hf(x,y) = \begin{bmatrix} 12x+10 & 2y \\ 2y & 2x+2 \end{bmatrix}$

$f(1,0) = 2 + 0 + 5 + 0 = 7$, $\nabla f(1,0) = \begin{bmatrix} 16 \\ 0 \end{bmatrix}$, $Hf(1,0) = \begin{bmatrix} 22 & 0 \\ 0 & 4 \end{bmatrix}$

$\vec a = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\vec h = \begin{pmatrix} h \\ k \end{pmatrix}$

$$f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \vec h\right) = f(1+h,\ 0+k) \approx 7 + \begin{bmatrix} 16 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} h \\ k \end{bmatrix} + \frac{1}{2} (h,k)^T Hf\begin{pmatrix} 1 \\ 0 \end{pmatrix} (h,k)$$

$$= 7 + 16h + \frac{1}{2}\left(22h^2 + 4k^2 + 2\cdot 0hk\right)$$

$$= \boxed{11h^2 + 2k^2 + 16h + 7}$$

b) crit pnts: $(0,0)$, $(-5/3, 0)$, $(-1,2)$, $(-1,-2)$

c) $\lambda = 12, 2$ local min; $\lambda =$ local max; saddle; saddle

d) $\vec w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$; ellipse; $\vec w = \begin{pmatrix} -\sqrt{17}-1 \\ 4 \end{pmatrix}$ & $\begin{pmatrix} \sqrt{17}-1 \\ 4 \end{pmatrix}$