(3.3)

"How heigh will the sycamore grow? If you cut it down, then you'll never know" sings Pocahontas as I and many hapless others were forced to sing in my heigh-school's music class. Unfortunately, mechanical analysis suggests that cutting the tree into cylindrical segments is not unreasonable as a procedure for determining maximal tree length (Which beggs the question: How high do music teachers grow?).

According to John A. Adam's "A mathematical nature walk", the size of a tree trunk can be approximated by a cylindrical beam of uniform density. The height of that beam, H, cannot exceede $\left(\frac{E}{\rho}\right)^{1/3} D^{2/3}$, where E is the stress-to-strain ratio, ρ is the density of beam and D is its diameter. It follows that $H_{max} = \left(\frac{E}{\rho}\right)^{1/3} D^{2/3}$.

Just like the problem proposed by Pocahontas, many important problems are problems of optimization. In this section, we develop procedures for solving optimization problems of several variables. In particular, we seek to identify critical points and classify them as local maximum or local minimum.

Def: A function $f: \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ has a local minimum at $a \in \mathcal{U}$ if there is an open ball B about a such that $f(x) \geq f(a)$ for all $x \in B$. Similarly, f has a local maximum at $a \in \mathcal{U}$ if $f(x) \leq f(a)$ for all x is some open ball about a. In either of these cases we say that f has a local extremum at a.

Ex. Find the critical points of

a) $1-\sqrt{x^2+y^2}$

b) $(x-2)^2+(y+1)^2+3$

c) $(x^2+y^2)^{3/2}-3(x^2+y^2)^{1/2}+5$

Solution:

a) $f(x,y)=1-\sqrt{x^2+y^2}$ is a surface of revolution of $g(x)=1-x$. Therefore the graph of f is a cone with vertex at $(0,0,1)$ that opens downward. Thus f has a local maximum at $(0,0)$ whose value is 1. Since 1 is the heighest point on the surface of f, f has a global maximum at $(0,0)$. Notice that f is not differentiable at $(0,0)$

b) Let $f(x,y)=(x-2)^2+(y+1)^2+3$. Then $f(x,y)\equiv u^2+v^2+3$, which is a paraboloid. Notice that $u^2+v^2+3\geq 3$ with equality holding when $u^2+v^2=0$. This happens when $x=2$ and $y=-1$. Thus f has a global minimum at $(2,-1)$. The value of this minimum is 3.

c) $f(x,y)=(x^2+y^2)^{3/2}-3(x^2+y^2)^{1/2}+5$ is a surface of revolution of $g(x)=x^3-3x+5$ $x\geq 0$. We can locate all the critical points of f by locating the corresponding critical points of g. Since g is a differentiable function of one variable, we can use techniques from single variable calculus: $g'(x)=3x^2-3=0$ when $x=\pm 1$. Since $x\geq 0$, the only critical points are $x=0$ and $x=1$. The number line for $g'(x)$ is $\begin{array}{ccccccc} + & 0 & - & & - & 0 & + \\ & -1 & & 0 & & 1 & \end{array}$
so $x=1$ is a local minimum and $x=0$ corresponds to a local maximum. At $x=0$, the height $g(0)=5$, while at $x=1$, $g(1)=3$.

It follows that f has a local maximum at $\sqrt{x^2+y^2}=0 \Rightarrow (x,y)=(0,0)$ and a local minimum at $\sqrt{x^2+y^2}=1 \Rightarrow x^2+y^2=1$. In other words, f has a local min at every point (x,y) on the unit circle.

Just like in single verriate calculus, the critical points of a function are points at which the function is not differentiable or at which the derivative is 0:

Thm: If $f: U \subset \mathbb{R}^n \to \mathbb{R}$ has a local extremum at a point $a \in U$, then either $\nabla f(a) = 0$ or the gradient at a is undefined.

Proof: First, suppose that f has a local maximum at $a = (a_1, a_2, \ldots, a_n)$. Then there exists an open ball B about a s.t. $f(a) \geq f(x)$ for all $x \in B$. Now the gradient is either defined at a or else it isn't. If $\nabla f(a)$ is defined, then $\frac{\partial f}{\partial x_i}(a)$ is defined for $i = 1, 2, \ldots, n$. Observe that $\frac{\partial f}{\partial x_i}(a)$ is the ordinary derivative of the function $g(x) = f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n)$ evaluated at $x_i = a_i$. Since f has a local maximum at a, we know g has a local maximum at a_i. Thus, by single-variable calculus, $g'(a_i) = 0$. But $g'(a_i) = \frac{\partial f}{\partial x_i}(a) = 0$. This reasoning applies for each $i = 1, 2, \ldots, n$ so $\nabla f(a) = 0$. The alternative in this case is that the gradient is undefined at a.

If f has a local minimum at a, the same reasoning again leads to one of two alternatives: $\nabla f(a) = 0$ or the gradient is undefined at a.

Def: A point a in the domain of $f: U \subset \mathbb{R}^n \to \mathbb{R}$ is called a critical point of f if either $\nabla f(a) = 0$ or the gradient of f is undefined at a.

With this definition we can restate the above theorem in a useful form.

Corollary: The only points in the domain of a function f at which it may have a local extremum are the critical points.

Critical points that correspond to neither a local maximum nor a local minimum are called saddle points.

Ex. Let $f(x,y,z) = x^4 + y^2 + z^3$. Then $\nabla f(x,y,z) = (4x^3, 2y, 3z^2)$. All the critical points of f correspond to solutions of the equation $\nabla f(x,y,z) = (0,0,0)$. It follows that $(0,0,0)$ is the only critical point. Observe that $(0,0,0)$ is neither a local max nor a local min. $(0,0,0)$ is not a local max because, for any $\delta > 0$, the point $(0,0\tfrac{1}{2}\delta) \in B_\delta(0,0,0)$ satisfies $f(0,0,\tfrac{1}{2}\delta) = \tfrac{1}{8}\delta^3 > 0 = f(0,0,0)$.

It is not a local min because, for any $\delta > 0$, the point $(0,0,-\tfrac{1}{2}\delta) \in$ $\in B_\delta(0,0,0)$ satisfies $f(0,0,-\tfrac{1}{2}\delta) = -\tfrac{1}{8}\delta^3 < 0 = f(0,0,0)$.

Hence f has a saddle point at $(0,0,0)$.

Ex. Classify the critical points of $f(x,y) = (x-1)^{2/3}(y+2)^{2/3}$.

Solution: $\nabla f(x,y) = \left(\frac{2}{3}(x-1)^{-1/3}(y+2)^{2/3}, \frac{2}{3}(x-1)^{2/3}(y+2)^{-1/3}\right)$.

The gradient is undefined whenever $x-1=0$ or $y+2=0$.

Thus, any point of the form $(1,y)$ or $(x,-2)$ are critical points.

To classify these points, observe that $f(x,y) \equiv u^{2/3}v^{2/3} = (uv)^{2/3} =$ $= \left(\frac{1}{2}\right)^{2/3}(2uv)^{2/3} \equiv \left(\frac{1}{2}\right)^{2/3}\left(2(r\cos\theta)(r\sin\theta)\right)^{2/3} = \left(\frac{r^2}{2}\right)^{2/3}\sin^{2/3}2\theta \geq 0$

When $\theta = 0, \frac{\pi}{2}, \pi$, and 2π $\left(\frac{r^2}{2}\right)^{2/3}\sin^{2/3}2\theta = 0$. But $\theta = 0, 2\pi$ corresponds to the u-axis (points of the form $(u,0)$) and $\theta = \frac{\pi}{2}$ corresponds to the v-axis (points of the form $(0,v)$). This means that $(u,0) \equiv$ $(x,-2)$ and $(0,v) \equiv (1,y)$ are local minima.

Ex. Find the local extrema of $f(x,y) = x^3 + x^2 y - y^2 - 4y$

Solution: $\nabla f(x,y) = (3x^2 + 2xy,\ x^2 - 2y - 4) = (0,0)$

$$3x^2 + 2xy = 0$$
$$x^2 - 2y - 4 = 0$$

The second equation implies that $2y = x^2 - 4$. Substituting into the first equation we get that $3x^2 + x(x^2-4) = 0 \Rightarrow$
$x(3x + x^2 - 4) = 0 \Rightarrow x(x-1)(x+4) = 0$

If $x = 0$ then $y = -2$. If $x = 1$ then $y = -\frac{3}{2}$. Finally, if $x = -4$, $y = 6$. Thus the critical points are $(0,-2)$, $(1,-\frac{3}{2})$, and $(-4,6)$.

Unfortunately, it is not simply a matter of rewriting $f(x,y)$ as we did in the previous example to see whether f has extrema at these points. We will now develop a tool for making this determination.

The second derivative test for functions of several variables

---

Def: A function $g: \mathbb{R}^n \to \mathbb{R}$ is called quadratic if it is of the form $g(h_1, \dots, h_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j$.

If A is the $n \times n$ matrix (a_{ij}), we can write $g(h_1, h_2, \dots, h_n)$ as

$$(h_1 \dots h_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = (h_1 \dots h_n) A \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} =$$

$= h A h^T$

Ex. Let $g(h_1, h_2, h_3, h_4) = 2h_1^2 - 4h_1h_2 + 7h_1h_3 + 3h_2^2 + 10h_2h_4 + h_4^2$.
Then g is a quadratic function of the form

$$\begin{pmatrix} h_1 & h_2 & h_3 & h_4 \end{pmatrix} \begin{pmatrix} 2 & -4 & 7 & 0 \\ 0 & 3 & 0 & 10 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{pmatrix}$$

(How did I get this matrix?)

We can, if we wish, assume that $A = (a_{ij})$ is symmetric; if $g(h_1, \dots, h_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j$, a_{ij} may be replaced by $\frac{1}{2}(a_{ij} + a_{ji})$ without changing the sum. This is because $a_{ij} h_i h_j + a_{ji} h_j h_i = (a_{ij} + a_{ji}) h_i h_j$. By letting $B = (b_{ij})$, where $b_{ij} = \frac{1}{2}(a_{ij} + a_{ji})$ we see that $(b_{ij})^T = (b_{ji}) = (b_{ij})$. Hence $B^T = B$ and B is symmetric. Furthermore, we can write
$g(h_1, \dots, h_n) = h B h^T$

Functions of the form $g(h_1, \dots, h_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j$ get their name, because $g(\lambda h_1, \dots, \lambda h_n) = \lambda^2 g(h_1, \dots, h_n)$ for any scalar $\lambda \in \mathbb{R}$.
A quadratic function $g : \mathbb{R}^n \to \mathbb{R}$ is called positive-definite if $g(h) \geq 0$ for all $h \in \mathbb{R}^n$ and $g(h) = 0$ if and only if $h = 0$. Similarly, g is negative-definite if $g(h) \leq 0$ and $g(h) = 0$ iff $h = 0$.

How does all of this help us classify critical points? Suppose $f: \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ is of class C^3 (i.e. f is 3 times continuously differentiable). Suppose further that f has a critical point $a \in \mathcal{U}$. In particular, $\nabla f(a) = 0$. By the work done in the previous chapter, $f(a+h) = f(a) + (h \cdot \nabla) f(a) + \frac{1}{2}(h \cdot \nabla)^2 f(a) + \mathcal{E}_2(a,h)$.
The error $\mathcal{E}_2(a,h) = \frac{1}{3!}(h\cdot\nabla)^3 f(x_0)$, where $x_0 = a + t_0 h$ for $t_0 \in (0,1)$.
Notice that the expression $\frac{1}{2}(h\cdot\nabla)^2 f(a) = \frac{1}{2}\left(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n}\right)^2 f(a)$ is a quadratic form $H(h_1,\dots,h_n) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, h_i h_j$

This quadratic function can be written as

$$\frac{1}{2}\begin{pmatrix} h_1 & \dots & h_n \end{pmatrix} \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

This quadratic function is called the Hessian of f at a. It is named after Ludwig Otto Hesse, who introduced it in 1844. We denote this function by $H(f(a))(h)$. With this notation, $f(a+h) = f(a) + (h\cdot\nabla)f(a) + H(f(a))(h) + \frac{1}{6}(h\cdot\nabla)^3 f(x_0)$.
We are finally ready to state and prove the second derivative test for local extrema.
Thm: If $f: \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ is of class C^3 and $a \in \mathcal{U}$ is a critical point of f, then

1) If the Hessian $Hf(a)$ is positive-definite, a is a relative min of f.

2) If the Hessian $Hf(a)$ is negative-definite, a is a relative max.

Proof: Suppose $Hf(a)$ is positive-definite. Then $Hf(a)(h) \geq 0$ for all $h \neq 0$. Observe that $Hf(a)$, being a quadratic form, is a polynomial. Therefore, $Hf(a)$ is continuous on $\mathbb{R}^n$. Let $K = \{h \in \mathcal{U} : \|h\| = 1\}$. Then K is the set of all unit vectors in $\mathcal{U}$. Notice that K is closed and bounded (Prove this!). By the extreeme-value theorem (see the chapter on limits), $Hf(a)$ attains a minimum, m, on K. In particular, for $h \in \mathcal{U}$ s.t. $h \neq 0$, $h = \frac{h}{\|h\|}\|h\|$ where $\frac{h}{\|h\|} \in K$, which implies that

$$Hf(a)(h) = Hf(a)\left(\frac{h}{\|h\|}\|h\|\right) = Hf(a)\left(\frac{h}{\|h\|}\right)\|h\|^2 \geq m\|h\|^2.$$

Observe that $|E_2(a,h)| = \left|\frac{1}{6}(\nabla \cdot h)^3 f(x_0)\right| = \left|\frac{1}{6}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x_0)h_i h_j h_k\right|$

$$\leq \frac{1}{6}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\left|\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x_0)\right| |h_i h_j h_k| \leq$$

$$\leq \frac{1}{6}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\left|\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x_0)\right| \|h\|^3 \text{ (Why?)} \qquad (1)$$

Because $\left|\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}\right|$ is continuous on $\mathcal{U}$ for any choice of i, j, k, it attains a maximum over any closed ball $C_\delta(0) = \{h \in \mathcal{U} : \|h\| \leq \delta\}$. Denote this maximum by M_{ijk}. Letting $M = \frac{1}{6}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} M_{ijk}$, we see that $(1) \leq M\|h\|^3$.

It follows that $Hf(a)(h) + \frac{1}{6}(h \cdot \nabla)^3 f(x_0) \geq m\|h\|^2 - M\|h\|^3$. Since

$$\lim_{h \to 0} \frac{m\|h\|^2 - M\|h\|^3}{\|h\|^3} = \lim_{h \to 0}\left(\frac{m}{\|h\|} - M\right) = \infty, \text{ it follows that}$$

for $\|h\| < \delta$ (for some δ) $\frac{m}{\|h\|} - M > 0 \Rightarrow m\|h\|^2 - M\|h\|^3 > 0$

Therefore, for $\|h\|$ small enough, $f(a+h) = f(a) + Hf(a)(h) + \frac{1}{6}(h\cdot\nabla)^3 f(x$
$\geq f(a) + m\|h\|^2 - M\|h\|^3 > f(a)$. This shows that a is a relative min.

The proof that a is a relative max when $Hf(a)$ is negative definite is similar.

It follows that classifying a critical point as a local max or local min may boil down to determining whether $Hf(a)$ is positive-definite or negative-definite. We state a criterion for determining this in the next theorem.

Thm: (second derivative test). Suppose $f: U \subset \mathbb{R}^n \to \mathbb{R}$ has continuous second order partials in an open ball B about a critical point $a \in U$ where $\nabla f(a) = 0$. Let $(h_{ij}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$. That is, (h_{ij}) is the Hessian matrix. Define $H_1 = (h_{11})$, $H_2 = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$, $H_3 = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}, \dots$

$$H_n = \begin{pmatrix} h_{11} & h_{12} & h_{13} & \dots & h_{1n} \\ h_{21} & h_{22} & h_{23} & \dots & h_{2n} \\ h_{31} & h_{32} & h_{33} & \dots & h_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ h_{n1} & h_{n2} & h_{n3} & \dots & h_{nn} \end{pmatrix}$$

Then.

1. If $\det(H_1) > 0$, $\det(H_2) > 0$, $\det(H_3) > 0, \dots \det(H_n) > 0$ then f has a local minimum at a.

2. If $\det(H_1) < 0$, $\det(H_2) > 0$, $\det(H_3) < 0, \ldots$ (i.e. if the signs of the determinants alternate starting with $\det(H_1)$ negative) then f has a local maximum at a.

3. If $\det(H_i) \neq 0$ for $i = 1, \ldots, n$, but the signs do not follow one of the patterns in 1 or 2, then f has a saddle at a.

4. If $\det(H_i) = 0$ for some i, then this test provides no information.

Proof: The proof is beyond the scope of this text. A proof can be found in "Shilov's Linear Algebra".

Ex. Recall that the critical points of $f(x,y) = x^3 + x^2y - y^2 - 4y$ are $(0,-2)$, $(1, -\frac{3}{2})$, $(-4,6)$

The Hessian matrix of f is $\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6x+2y & 2x \\ 2x & -2 \end{pmatrix}$

Thus $H(0,-2) = \begin{pmatrix} -4 & 0 \\ 0 & -2 \end{pmatrix}$ and $\det(-4) < 0$, $\det\begin{pmatrix} -4 & 0 \\ 0 & -2 \end{pmatrix} > 0$.

Thus f has a local max at $(0,-2)$.

$H(1, -\frac{3}{2}) = \begin{pmatrix} 3 & 2 \\ 2 & -2 \end{pmatrix}$ and $\det(3) > 0$, $\det\begin{pmatrix} 3 & 2 \\ 2 & -2 \end{pmatrix} = -10 < 0$

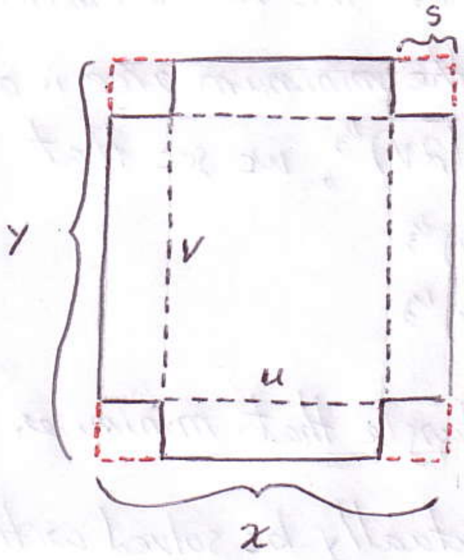
So $(1, -\frac{3}{2})$ is a saddle point of f.

Finally $H(-4,6) = \begin{pmatrix} -12 & -8 \\ -8 & -2 \end{pmatrix}$ and $\det(-12) < 0$, $\det\begin{pmatrix} -12 & -8 \\ -8 & -2 \end{pmatrix} = -40 < 0$

So $(-4,6)$ is a saddle point of f.

Ex. (From Thomas Barr' "Vector Calculus")

An open-top rectangular box of a specified volume V is to be constructed from sheet metal by first cutting a rectangular piece of sheet, then cutting squares from the corners, and finally folding up the remaining flaps and soldering their ends together. Find the dimensions of the material that minimize the amount of sheet metal used.



Solution: We want to find x, y, and s to minimize the area

$A = (x-2s)(y-2s) + 2(x-2s)s + 2(y-2s)s$,

where the variables satisfy $(x-2s)(y-2s)s = V$. To simplify matters, we first substitute $u = x-2s$ and $v = y-2s$ so that

$A = uv + 2us + 2vs = uv + 2s(u+v)$ is subject to the constraint $V = uvs$.

This means that we can replace s by $\frac{V}{uv}$.

Therefore, we want to minimize the function $A(u,v) = uv + \frac{2V}{uv}(u+v) =$

$= uv + 2V\left(\frac{1}{v} + \frac{1}{u}\right)$, $u, v > 0$

$\nabla A = \left(v - \frac{2V}{u^2}, u - \frac{2V}{v^2}\right) = (0,0)$ when $v = \frac{2V}{u^2}$ and $u = \frac{2V}{v^2}$

or, upon substituting $\frac{2V}{u^2}$ in place of v, $u = \frac{2V}{(2V/u^2)^2} =$
$= 2V \cdot \frac{u^4}{(2V)^2} = \frac{u^4}{2V}$. This can be simplified to $u^4 - 2Vu = 0$
Since $u > 0$, this is further simplified to $u^3 = 2V$ or $u = (2V)^{1/3}$.
Substituting $u = (2V)^{1/3}$ in place of $v = \frac{2V}{u^2}$ shows that $v = (2V)^{1/3}$ as well. It follows that $((2V)^{1/3}, (2V)^{1/3})$ is a critical point of $A(u,v)$. You should verify that the Hessian matrix is positive definite at this point. Thus, the minimum area is attained for $u = v = (2V)^{1/3}$. Since $s = \frac{V}{uv} = \frac{1}{2}(2V)^{1/3}$, we see that

$$x = u + 2s = 2(2V)^{1/3}$$
$$y = v + 2s = 2(2V)^{1/3}$$

are the dimensions of the rectangle that minimizes the area.

Remark: This problem can actually be solved with techniques from single-variable Calculus. To see this, observe that $A(u,v) =$
$= uv + 2V\left(\frac{1}{v} + \frac{1}{u}\right)$ is a symmetric function. If we reverse the variables, writing $A(v,u)$ instead of $A(u,v)$, the function is unchanged. This means that if (a,b) is a critical point of A then so is (b,a). Observe that
$\frac{\partial A}{\partial u} = v - \frac{2V}{u^2} = 0 \Rightarrow \frac{2V}{u^2} = v$ since $f(u) = \frac{2V}{u^2}$ is 1-1 for $u > 0$
there is only 1 critical point. It follows that A has only 1 critical point (a,b). Since, by symmetry of A, (b,a) is also a critical point, it follows that $(a,b) = (b,a)$ or $a = b$.
Thus we may search for the critical points when $u = v$; $A(u,v)\big|_{u=v} =$
$= u^2 + \frac{4V}{u} = F(u)$. Now $\frac{dF}{du} = 2u - \frac{4V}{u^2} = 0$ for $u = (2V)^{1/3}$

Optimization techniques can be used to find polynomial approximations to non-polynomial curves.

Ex. Find the second-degree polynomial $p(t) = x + yt + zt^2$ that minimizes $\int_0^1 \left(p(t) - \frac{1}{t+1}\right)^2 dt = f(x,y,z)$.

The function f is the best least squares fit of a second-degree polynomial to the rational function $1/_{t+1}$ on the interval $[0,1]$.

Solution: Upon integrating, we see that $f(x,y,z) = x^2 + \frac{1}{3}y^2 + \frac{1}{5}z^2 + xy +$
$+ \frac{1}{2}yz + \frac{2}{3}xz - (2\ln 2)x + 2(\ln 2 - 1)y + (1 - 2\ln 2)z + \frac{1}{2}$.

$\nabla f(\vec{x}) = \left(2x + y + \frac{2}{3}z - 2\ln 2,\ x + \frac{2}{3}y + \frac{1}{2}z + 2(\ln 2 - 1),\ \frac{2}{3}x + \frac{1}{2}y + \frac{2}{5}z + (1 - 2\ln 2)\right)$

The gradient $\nabla f(\vec{x}) = 0$ when

$$\begin{pmatrix} 2 & 1 & \frac{2}{3} \\ 1 & \frac{2}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} & \frac{2}{5} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2\ln 2 \\ -2(\ln 2 - 1) \\ -(1 - 2\ln 2) \end{pmatrix}$$

By using Cramer's rule (see p. 42 in your book or optional reading in my notes on Linear Algebra), we see that the only solution is

$$\vec{a} = (-51 + 75\ln 2,\ 282 - 408\ln 2,\ -270 + 390\ln 2) \approx (0.9860, -0.8041, 0.3274)$$

Since the Hessian matrix of f, $\begin{pmatrix} 2 & 1 & \frac{2}{3} \\ 1 & \frac{2}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} & \frac{2}{5} \end{pmatrix}$, is positive definite,

we see that the best least squares fit is the polynomial

$$p(t) = (-51 + 75\ln 2) + (282 - 408\ln 2)t + (-270 + 390\ln 2)t^2$$

or $0.9860 - 0.8041t + 0.3274t^2$

Ex. The function $f(x,y) = x^4 + y^4$ has a single local minimum at $(0,0)$. However, $H_f(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Thus $\det(H_1) = 0 = \det(H_2)$. This shows that in the situation when two of the determinants of the submatrices are 0, there may be a local minimum.
The function $g(x,y) = x^3 + y^3$ does not have a local minimum at its critical point $(0,0)$. Here $H_g(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
Again $\det(H_1) = \det(H_2) = 0$. This illustrates why $\det(H_i) = 0$ provides no information about the critical point.

In single-variable calculus, you learned the following fact: If a function f has first and second-order derivatives that are continuous at a number a, and if $f'(a) = 0$ and $f''(a) > 0$, then f has a local minimum at a. With this in mind, consider a function f of two variables with first- and second-order partials that are continuous at a point (a_1, a_2). Suppose $\nabla f(a) = 0$ and that every vertical plane through the point $(a_1, a_2, f(a_1, a_2))$ intersects the graph of f in a curve that is concave upward. We would say with virtual certainty that f has a local minimum at (a_1, a_2). Yet this is not necessarily true! Consider the following example.
Ex. (Thomas Barr)
Rebecca has been hiking in the mountains and valleys of eastern Tennessee, and she stops to rest. Looking straight ahead, she notes that the ground slopes upward. Rotating her head just a bit, she sees that the same is true in that direction. Slowly, she turns all the way around at her resting spot and she sees that in every direction, the ground appears to slope upward. Rebecca must be at the bottom of a valley, right? Wrong.

Suppose that the landscape is described by the graph of

$$z=(y-8x^2)(y-x^2)$$

and that she sits at the point $(0,0,0)$. Along any line $y=mx$, $m\neq 0$, through the origin in the xy-plane

$$z=(mx-8x^2)(mx-x^2)=m^2x^2-9mx^3+8x^4 \quad \text{so}$$

$$\frac{dz}{dx}=2m^2x-27mx^2+32x^3 \quad \text{and}$$

$$\frac{d^2z}{dx^2}=2m^2-54mx+96x^2.$$

When $x=0$, $\frac{dz}{dx}=0$ and $\frac{d^2z}{dx^2}=2m^2>0$, so along the line $y=mx$, the graph has a strict local minimum. (You should also verify that even along the lines $y=0$ and $x=0$, the function has a local minimum.) This function, then, exhibits exactly the shape that Rebecca is observing.

But notice what happens if she walks away from $(0,0,0)$ along a curve on the surface corresponding to $y=5x^2$. Her altitude is given by $z=(5x^2-8x^2)(5x^2-x^2)=-12x^2$; she will go downhill.

How can this be? Philosophically, this means that curves are qualitatively different from lines. In particular, just like irrational numbers cannot be described by a finite sequence of decimals, the curve $y=x^2$ cannot be described as consisting of finitely many line segments. Thus, moving along a parabola is fundamentally different from moving along a line. (Does this conflict with the idea that every curve may be thought of as a line in a curved space?)