Quadratic Forms
Our study of extrema of multivariable functions has focused so far on:

- Conditions on the function necessary for a local or global extremum to occur \((\nabla f(\vec{a}) = \vec{0}, \ \vec{a} = \text{boundary/limit point})\)
- Conditions on the domain \(S\) that ensure any continuous function has global extrema \((S\text{ is closed + bounded, i.e. compact})\)

Today we will discuss how the second partials of a \(C^2\) function tell us when a stationary point \((\nabla f(\vec{a}) = \vec{0})\) is a local maximum, local minimum, or saddle point. This is one of the first places matrix algebra really "barges in" to Calculus.

\[8.1. \text{The 2nd Derivative Test}\]

In the first example of Lecture 12, we looked at \(f(y) = x^2 + 4xy + y^2\) near \(\vec{a} = (0)\). The surprise was that, although \(f_{xx}(0) = \frac{\partial^2 f}{\partial x^2}(0) = 2 = f_{yy}(0)\), we actually had a saddle point at \(\vec{a}\).

We detected this by computing (with \(\vec{u} = \frac{1}{\sqrt{2}} \vec{i}\))

\[D^2_{\vec{u}}f(\vec{0}) = D_{\vec{u}}\{f(x^2 - y^2)\}(\vec{0}) = -2.\]

Note: Even weirder things can happen. \(f(y) = 2x^4 - 3x^2y + y^2\) has local minimum at \(\vec{0}\) along all lines through the origin, but still has a saddle point there! (Consider its restriction to \(y = \frac{3}{2}x^2\), which has a local MAX at \(\vec{0}\).) This funny business is only possible, though, because one of the \(D^2_{\vec{u}}f(\vec{0})\)'s is \(0\). If they're all positive, we'll have a local minimum.
To explore further, let $\hat{a}$ be a stationary point of a function $f(x,y)$, and set $\hat{u} = \left(\cos \theta \right) u_x + \left(\sin \theta \right) u_y$.

$$D^c_{\hat{u}} f = D_{\hat{u}} (u_x f_x + u_y f_y)$$
$$= \nabla (u_x f_x + u_y f_y) \cdot \hat{u}$$
$$= (u_x f_{xx} + u_y f_{xy}, u_x f_{xy} + u_y f_{yy}) \cdot (u_x, u_y)$$
$$= (u_x, u_y) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix}$$
$$= t_{\hat{u}} H \hat{u} \ (= \hat{u} \cdot H \hat{u})$$

So $\hat{u} \cdot H \hat{u}$ is the concavity at $\hat{a}$ in the direction $\hat{u}$, which depends on $\theta$. We want the directions of extreme concavity.

The $2 \times 2$-matrix-valued function $H(x,y)$ is called the Hessian of $f$. Write $H_0$ for $H(\hat{a})$. The unit circle is compact, so a max/min exist.
To explore further, let \( \hat{a} \) be a stationary point of a function \( f(x,y) \), and set \( \hat{u} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \) and \( u_y \).

\[
D_{\hat{u}} f = D\hat{u} (u_x f_x + u_y f_y) \\
= \nabla (u_x f_x + u_y f_y) \cdot \hat{u} \\
= (u_x f_{xx} + u_y f_{yx}, u_x f_{xy} + u_y f_{yy}) \cdot (u_x, u_y) \\
= (u_x, u_y) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix} \\
= u_x H \hat{u} \hat{u} (= \hat{u} \cdot H \hat{u})
\]

The 2x2 - matrix - valued function \( H(x,y) \) is called the Hessian of \( f \). Write \( H_0 \) for \( H(\hat{a}) \).
To explore further, let \( \hat{a} \) be a stationary point of a function \( f(x, y) \), and set \( \hat{u} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \).

\[
D_{\hat{u}}f = D_{\hat{u}} (u_x f_x + u_y f_y)
= \nabla (u_x f_x + u_y f_y) \cdot \hat{u}
= (u_x f_{xx} + u_y f_{xy}, u_x f_{xy} + u_y f_{yy}) \cdot (u_x, u_y)
= (u_x, u_y) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix}
= t^\top \hat{u} H \hat{u} (= \hat{u} \cdot H \hat{u})
\]

So \( \hat{u} \cdot H \hat{u} \) is the concavity at \( \hat{a} \) in the direction \( \hat{u} \), which depends on \( \theta \).

We want the directions of extreme concavity:

\[
0 = \frac{d}{d\theta} \hat{u} \cdot H \hat{u} = \hat{u} \cdot H \hat{u} + \hat{u} \cdot H \hat{u}^	op
\]

\( \Rightarrow 2 \hat{u} \cdot H \hat{u} \Rightarrow \hat{u} \perp H \hat{u} \)

The \( 2 \times 2 \) matrix-valued function \( H(x, y) \) is called the Hessian of \( f \).

Write \( H_0 \) for \( H(\hat{a}) \).
To explore further, let \( \hat{a} \) be a stationary point of a function \( f(x,y) \), and set \( \hat{u} = (\cos \theta) u_x + (\sin \theta) u_y \)

\[
D^c_{\hat{u}} f = D_{\hat{u}} (u_x f_x + u_y f_y) \\
= \nabla (u_x f_x + u_y f_y) \cdot \hat{u} \\
= (u_x f_{xx} + u_y f_{xy}, u_x f_{xy} + u_y f_{yy}) \cdot (u_x, u_y) \\
= (u_x, u_y) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix} \\
= \hat{u} \cdot H \hat{u} \quad (= \hat{u} \cdot H \hat{u})
\]

So \( \hat{u} \cdot H_0 \hat{u} \) is the concavity at \( \hat{a} \) in the direction \( \hat{u} \), which depends on \( \theta \).

We want the directions of extreme concavity:

\[
0 = \frac{d}{d\theta} \hat{u} \cdot H_0 \hat{u} = \hat{u} \cdot H_0 \hat{u} + \hat{u} \cdot H_0 \hat{u} \cdot \hat{u}'
\]

\[
= 2 \hat{u}' \cdot H_0 \hat{u} \quad \Rightarrow \quad \hat{u}' \perp H_0 \hat{u}
\]

\[\Rightarrow \quad \hat{u} \parallel H_0 \hat{u} \quad \Rightarrow \quad \hat{u} \text{ eigenvector of } H_0 !
\]

\[\hat{u}' = \frac{d}{d\theta} (\cos \theta) = (\cos \theta)
\]

is \( \perp \) to \( \hat{u} \)

The \( 2 \times 2 \) matrix-valued function \( H(x,y) \) is called the Hessian of \( f \). Write \( H_0 \) for \( H(\hat{a}) \).
To explore further, let \( \hat{a} \) be a stationary point of a function \( f(x,y) \), and set \( \hat{u} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \). 

\[
D_{\hat{u}}^c f = D\hat{u} (u_x f_x + u_y f_y) \\
= \nabla (u_x f_x + u_y f_y) \cdot \hat{u} \\
= (u_x f_{xx} + u_y f_{xy}, u_x f_{xy} + u_y f_{yy}) \cdot (u_x, u_y) \\
= (u_x, u_y) \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix} \\
= \hat{u}^T H \hat{u} \quad (= \hat{u} \cdot H \hat{u})
\]

The 2x2 matrix-valued function \( H(x,y) \) is called the Hessian of \( f \).

Write \( H_0 \) for \( H(\hat{a}) \).

So \( \hat{u} \cdot H_0 \hat{u} \) is the concavity at \( \hat{a} \) in the direction \( \hat{u} \), which depends on \( \theta \).

We want the directions of extreme concavity:

\[
0 = \frac{d}{d\theta} \hat{u} \cdot H_0 \hat{u} = \hat{u} \cdot H_0 \hat{u} + \hat{u} \cdot H_0 \hat{u}' \\
= 2 \hat{u} \cdot H_0 \hat{u} \quad \Rightarrow \quad \hat{u}' \perp H_0 \hat{u} \\
\quad \Rightarrow \quad \hat{u} \parallel H_0 \hat{u} \quad \Rightarrow \quad \hat{u} \text{ eigenvector of } H_0
\]

If \( \hat{u} = \nabla \lambda = \text{eigenvector w/eigenvalue } \lambda \) then \( D_{\hat{u}}^c f = \nabla \cdot H_0 \nabla \lambda = \nabla \lambda \parallel \hat{u} \Rightarrow \lambda \parallel \hat{u} \quad \Rightarrow \quad \lambda \text{ is the concavity.}
\]

Writing \( \lambda_1 \leq \lambda_2 \) for the 2 eigenvalues, these are the min/max. concavities of \( f \) at \( \hat{a} \).

**FACT 1** \( \Delta := \det \begin{pmatrix} f_{xx}(\hat{a}) & f_{xy}(\hat{a}) \\ f_{yx}(\hat{a}) & f_{yy}(\hat{a}) \end{pmatrix} \) is the product of the max/min concavities of \( f \) at \( \hat{a} \).
But if $\Delta > 0$ this tells me only that $\lambda_1, \lambda_2 > 0 \Rightarrow \lambda_1, \lambda_2 < 0$. Which is it? Consider $f_{xx}(\hat{a})$, the concavity in the $x$-direction—such, clearly in between the max/min concavities: $\lambda_1 \leq f_{xx}(\hat{a}) \leq \lambda_2$.

So if $f_{xx}(\hat{a}) > 0$, then $\lambda_2 > 0$; and since we assumed $\Delta = \lambda_1 \lambda_2 > 0$, $\lambda_1 > 0$ too! $\Rightarrow$

**FACT 2**: If $\Delta > 0$ and $f_{xx}(\hat{a}) > 0$ then $\lambda_1, \lambda_2 > 0$ and $\hat{a}$ is a local minimum.

- If $\Delta > 0$ and $f_{xx}(\hat{a}) < 0$ then $\lambda_1, \lambda_2 < 0$ and $\hat{a}$ is a local maximum.
- If $\Delta < 0$, then $\lambda_1 < 0, \lambda_2 > 0$, and $\hat{a}$ is a saddle point.
- If $\Delta = 0$, the test is inconclusive and you need to try something else!

So $\hat{u} \cdot H_0 \hat{u}$ is the concavity at $\hat{a}$ in the direction $\hat{u}$, which depends on $\Theta$. We want the directions of extreme concavity:

$$0 = \frac{d}{d\Theta} \hat{u} \cdot H_0 \hat{u} = \hat{u} \cdot H_0 \hat{u} + \hat{u} \cdot H_0 \hat{u}'$$
$$= 2 \hat{u}' \cdot H_0 \hat{u} \Rightarrow \hat{u}' \perp H_0 \hat{u} \Rightarrow \hat{u} \parallel H_0 \hat{u} \Rightarrow \hat{u} \text{ eigenvector of } H_0.$$

If $\hat{u} = \hat{v}$ is eigenvector w/eigenvalue $\lambda$, then $D^2_\hat{u} f = \hat{v} \cdot H_0 \hat{v} = \hat{v} \cdot \lambda \hat{v} = \lambda \| \hat{v}^2 \Rightarrow$ $\lambda$ is the concavity.

Writing $\lambda_1 \leq \lambda_2$ for the 2 eigenvalues, these are the min/max. concavities of $f$ at $\hat{a} \Rightarrow$

**FACT 1**: $\Delta := \det \begin{pmatrix} f_{xx}(\hat{a}) & f_{xy}(\hat{a}) \\ f_{yx}(\hat{a}) & f_{yy}(\hat{a}) \end{pmatrix}$ is the product of the max/min concavities of $f$ at $\hat{a}$. 
Given \( f \) twice continuously differentiable on a neighborhood of \( \hat{a} \), with \( \nabla f(\hat{a}) = 0 \).

**FACT 1** \( \Delta := \det \begin{pmatrix} f_{xx}(\hat{a}) & f_{xy}(\hat{a}) \\ f_{yx}(\hat{a}) & f_{yy}(\hat{a}) \end{pmatrix} \) is the product of the max/min concavities of \( f \) at \( \hat{a} \).

**FACT 2**
- If \( \Delta > 0 \) and \( f_{xx}(\hat{a}) > 0 \) then \( \lambda_1, \lambda_2 > 0 \) and \( \hat{a} \) is a local minimum.
- If \( \Delta > 0 \) and \( f_{xx}(\hat{a}) < 0 \) then \( \lambda_1, \lambda_2 < 0 \) and \( \hat{a} \) is a local maximum.
- If \( \Delta < 0 \), then \( \lambda_1 < 0, \lambda_2 > 0 \), and \( \hat{a} \) is a saddle point.
- If \( \Delta = 0 \), the test is inconclusive and you need to try something else!

Remark: These correspond to the quadratic form \( Q(\mathbf{h}) = \mathbf{h}^T \mathbf{H}(\hat{a}) \mathbf{h} \) being positive definite/negative definite/indefinite.

Returning to the Example: \( \hat{a} = \vec{0}, \)
\[ f(x, y) = x^2 + 4xy + y^2, \]
we have for the Hessian
\[ \mathbf{H}(x, y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix} \]
hence \( \Delta = 2^2 - 4^2 = -12 < 0 \)

\( \Rightarrow \) saddle.

Here are a couple of problems for you to try: find the local extrema and saddle points of

1. \( f(x, y) = x^4 + y^4 - 4xy + 1 \)

2. \( f(x, y) = x^8 - 3xy^2 \)
2. \( f(x, y) = x^8 - 3xy^2 \)

First solve \( f_x = 0 = f_y \rightarrow 8x^7 - 3y^2 = 0 \)
\[-6xy = 0 \]
\( \Rightarrow (x, y) = (0, 0) \) (why?).
So \( \hat{a} = (0, 0) \) is the only stationary point.
But \( \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 6x & -6y \\ -6y & -6x \end{pmatrix} \) is 0 at \( \hat{a} \),
hence \( \Delta = 0 \) so test in conclusive.
Rather, it's enough to write \( f(x, 0) = x^8 \)
to see that it's a saddle point — even in the \( x \)-dir. we don't have a max or min.
You could also go into greater detail by writing \( f(x, y) = x(x + \sqrt{3}y)(x - \sqrt{3}y) \).

\[ \begin{array}{cc}
+ & - \\
- & + \\
\hline
+ & - \\
\hline
x = 0 & x = \sqrt{3}y
\end{array} \]

and observing that the sign (of \( f(x, y) - f(x, 0) \)) changes as you go from chamber to chamber.

1. \( f(x, y) = x^4 + y^4 - 4xy + 1 \)

First solve \( f_x = 0 = f_y \rightarrow 4x^3 - 4y = 0 \)
\(-4y^3 + 4x = 0 \)
\( \Rightarrow x^2 = y \rightarrow (y^3)^2 = y \rightarrow 0 = y^6 - y = y(y^5 - 1) \)
\( \Rightarrow y = 0, 1, -1 \) (since y real) \( \Rightarrow \) the 3 stationary points are \( \hat{a} = (0, 0), (1, 1), (-1, -1) \).
Now \( H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 12x^2 - 4 \\ -4 & 12y^2 \end{pmatrix} \)
\( \Rightarrow \det H = 144x^2y^2 - 16 \). So

- at \( \hat{e} = (0, 0) \), \( \Delta = -16 < 0 \) so saddle pt.
- at \( \hat{e} = (1, 1) \), \( \Delta = 128 > 0 \) so local minimum \( \begin{cases} f_{xx} = 12 > 0 \\ f_{xy} = 12 > 0 \end{cases} \)
- at \( \hat{e} = (-1, -1) \), \( \Delta = 128 > 0 \) so local minimum \( \begin{cases} f_{xx} = 12 > 0 \\ f_{xy} = 12 > 0 \end{cases} \)
2. Eigenstuff (Cliff's Notes version)

Remembering a few "309" basics here really helps to understand what is going on in Section 2.5.3. Recall that for an $n \times n$ matrix $A$, an $n$-vector $\mathbf{v} \in \mathbb{R}^n$ is an eigenvector of $A$ with eigenvalue $\lambda \in \mathbb{R}$ if

$$A \mathbf{v} = \lambda \mathbf{v}.$$

- $A \mathbf{v} = \lambda \mathbf{v} \iff \mathbf{0} = (\lambda \mathbf{I} - A) \mathbf{v}$
  $\iff \mathbf{v} \in \text{null} (\lambda \mathbf{I} - A)$

- $\lambda$ is an eigenvalue $\iff \lambda \mathbf{I} - A$ is not invertible
  $\iff \det (\lambda \mathbf{I} - A) = 0$
  $\iff f_A(\lambda) = \det (\lambda \mathbf{I} - A)$

If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of $\mathbb{R}^n$ with $A \mathbf{v}_i = \lambda_i \mathbf{v}_i$ (an eigenbasis), set

$$S = (\mathbf{v}_1 \cdots \mathbf{v}_n).$$

so

$$S^{-1} A S = \begin{pmatrix} 0 & \cdots & \lambda_1 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} = D,$$

$\Rightarrow A = S D S^{-1}$.

We'll prove later that $\det (B C) = (\det B)(\det C)$
$\det (B^T) = 1/(\det B),$ so

- $\det (A) = \det (D) = \lambda_1 \cdots \lambda_n.$

If the $\{\mathbf{v}_i\}$ are orthonormal ($\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij}$) then

$$S S^T = (\begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}) = \begin{pmatrix} 1 & 0 \\ \vdots & 1 \end{pmatrix},$$

so

$$S^{-1} = S^T \Rightarrow \mathbf{y} = S \mathbf{x}.$$

- $Q(\mathbf{x}) := \mathbf{x}^T A \mathbf{x} = \mathbf{x}^T S D S^T \mathbf{x} = \mathbf{y}^T \mathbf{D} \mathbf{y}$
  $= (y_1^2 \cdots y_n^2)(\lambda_1 \cdots \lambda_n)(\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix})$
  $= \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$

Conclude that (in this case) if all $\lambda_i > 0$, then

$Q(\mathbf{x}) > 0$ for any $\mathbf{x} \neq \mathbf{0} \Rightarrow$ has minimum at $\mathbf{0}$. 

\[
\begin{align*}
\text{Conclude that (in this case) if all } &\lambda_i > 0, \text{ then } \\
Q(\mathbf{x}) &> 0 \text{ for any } \mathbf{x} \neq \mathbf{0} \Rightarrow \text{ has minimum at } \mathbf{0}.
\end{align*}
\]
Now suppose $A$ is a symmetric matrix. Then writing $Q(x) = A \cdot x$, 
$$
(DQ)_x h = (DAx)_x \cdot h + A \cdot (Dx)_x h \\
= 2A \cdot x \cdot h \Rightarrow \nabla Q = 2Ax
$$

Since the unit sphere $g(x) = \|x\|^2 = 1$ is compact, $Q(x)$ has a maximum at some $\hat{x}$. Finding $\hat{x}$ is a constrained extremum problem, which you learned to solve by using Lagrange multipliers, i.e. writing
$$
\nabla Q(\hat{x}) = \lambda \nabla g(\hat{x})
$$
and solving for $\lambda$ and $\hat{x}$:
$$
\begin{bmatrix}
2x_1 \\
2x_2 \\
\vdots \\
2x_n
\end{bmatrix} = \begin{bmatrix}
\hat{x}_1 \\
\hat{x}_2 \\
\vdots \\
\hat{x}_n
\end{bmatrix}
$$
$$
\Rightarrow 2Ax_1 = \lambda \hat{x}_1.
$$

So we have found an eigenvector (of unit length)!

Observe that for any $\hat{w} \perp \hat{x}$, 
$$
\hat{v} \cdot A \hat{w} = A \hat{v} \cdot \hat{w} = \lambda_1 \hat{v} \cdot \hat{w} = 0 \Rightarrow A \hat{w} \perp \hat{v}.
$$

Now the unit sphere intersected with the orthogonal complement 
$$
W := \{ f(x) = 0 \} \quad (f(x) := 2\hat{v} \cdot x)
$$

of $\hat{v}$ is compact, so again $Q(x)$ has a maximum at some $\hat{x}_2$. Since we have 2 constraints, we use 2 Lagrange multipliers:
$$
\nabla Q(\hat{x}_2) = \lambda_2 \nabla g(\hat{x}_2) + \mu_2 \nabla f(\hat{x}_2)
$$
$$
\frac{\nabla g(\hat{x}_2)}{2A \hat{v}_2} = \frac{\nabla f(\hat{x}_2)}{2\hat{v}_2}
$$
$$
A \hat{v}_2 = \lambda_2 \hat{v}_2 + \mu_2 \hat{v}_1
$$

and dottiing both sides with $\hat{v}_1$ gives $0 = \mu_2$ hence
$$
A \hat{v}_2 = \lambda_2 \hat{v}_2.
$$

Continuing in this fashion, we get an orthonormal eigenbasis of $\mathbb{R}^n$ (basis blc independent blc orthogonal - cf. HWS !).
Spectral Theorem: Given an \( n \times n \) symmetric matrix \( A \), there is an orthonormal basis \( \vec{v}_1, \ldots, \vec{v}_n \) of \( \mathbb{R}^n \) with \( A \vec{v}_i = \lambda_i \vec{v}_i \). Moreover, the associated quadratic form \( Q(\vec{x}) = \vec{x}^T A \vec{x} \) can be rewritten as \( \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \) in the coordinates \( \vec{y} = \sqrt{\lambda} \vec{x} \).

**Definition:**
(i) \( Q \) is

- positive definite \( \iff Q(\vec{x}) > 0 \ \forall \vec{x} \neq \vec{0} \)
- positive semidefinite \( \iff Q(\vec{x}) \geq 0 \ \forall \vec{x} \)
- negative definite \( \iff Q(\vec{x}) < 0 \ \forall \vec{x} \neq \vec{0} \)
- negative semidefinite \( \iff Q(\vec{x}) \leq 0 \ \forall \vec{x} \)
- indefinite \( \iff Q(\vec{x}) \) takes both positive and negative values

(ii) \( A \) is

- positive definite \( \iff \) all \( \lambda_i > 0 \)
- positive semidefinite \( \iff \) all \( \lambda_i \geq 0 \)
- negative definite \( \iff \) all \( \lambda_i < 0 \)
- negative semidefinite \( \iff \) all \( \lambda_i \leq 0 \)
- indefinite \( \iff \) \( A \) has both positive & negative eigenvalues.

**Corollary:** \( Q \) is \( \square \) \( \iff \) \( A \) is \( \square \).
Now we generalize to $n$ variables and examine $f(\xi)$ on the grey axis in the picture:

$$g(u) = f(a + uh) \text{ for } u \in [-1,1]$$

$$g''(u) = \frac{\partial^2 f}{\partial x_j \partial x_i} (a + uh)$$

By the chain rule again,

$$g''(u) = \left( \sum_{i,j} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j} (a + uh) \right)$$

$$= \frac{\partial}{\partial h} H(a + uh)$$

Tayloring remainder, $(0,1)$, Hessian of $f$.

$$f(a + h) - f(a) = g(1) - g(0) = g'(c) + \frac{1}{2} g''(c)$$

$$= \nabla f(a) \cdot h + \frac{1}{2} h^T H(a + ch) h$$

$$= \nabla f(a) \cdot h + \frac{1}{2} h^T H(a) h$$

$$+ O(h^2)$$

where $E_2(\hat{a}; \hat{h})$ is $o(1) \to 0 \text{ as } \hat{h} \to 0$.

assuming $f$ is twice continuously differentiable.

$\triangledown 3$. The 2nd Derivative Test (v.2)

$$\text{Where } E_2(\hat{a}; \hat{h}) := \begin{cases} \frac{1}{2} h^T H(a + ch) h & , \hat{h} \neq 0 \\ 0 & , \hat{h} = 0 \end{cases}$$

has

$$|E_2(\hat{a}; \hat{h})| = \frac{1}{2 \|h\|^2} \sum_{i,j} h_i h_j \left| \frac{\partial^2 f}{\partial x_i \partial x_j} (a + ch) - \frac{\partial^2 f}{\partial x_i \partial x_j} (a) \right|$$

$$\leq \frac{1}{2 \|h\|^2} \sum_{i,j} \|h_i h_j\| \left| \frac{\partial^2 f}{\partial x_i \partial x_j} (a + ch) - \frac{\partial^2 f}{\partial x_i \partial x_j} (a) \right|$$

$$\leq \frac{\|h\|^2}{2 \|h\|^2} \sum_{i,j} \left| \frac{\partial^2 f}{\partial x_i \partial x_j} (a + ch) - \frac{\partial^2 f}{\partial x_i \partial x_j} (a) \right|$$

$\to 0$ as $\hat{h} \to 0$ by continuity of 2nd partials.
Now we generalize to $n$ variables and examine $f(\hat{x})$ on the grey axis in the picture:

$g(u) = f(\hat{x} + uh)$ for $u \in [-1,1]$.

$g'(u) = \nabla f(\hat{x} + uh) \cdot h = \sum h_j \frac{\partial f}{\partial x_j}(\hat{x} + uh)$

Chain rule again:

$g''(u) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\hat{x} + uh) \cdot h = \sum h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\hat{x} + uh)$

$= h^T H(\hat{x} + uh) h = \left[ \frac{\partial^2 f}{\partial x_i \partial x_j}(\hat{x}) \right] h^T h$ (Taylor remainder)

Hessian of $f$

$\left. \begin{array}{c}
\nabla f(\hat{x}) \cdot h + \frac{1}{2} h^T H(\hat{x}) h
\n\end{array} \right]$

Theorem 2: If $f$ is twice continuously differentiable, with a stationary point at $\hat{x} = \hat{a}$, then:

(i) $H(\hat{a})$ pos.-definite $\Rightarrow f$ has rel. minimum at $\hat{a}$

(ii) $H(\hat{a})$ neg.-definite $\Rightarrow f$ has rel. maximum at $\hat{a}$

(iii) $H(\hat{a})$ indefinite $\Rightarrow f$ has saddle point at $\hat{a}$.

Proof: We record (before erasing) the fact that

$\nabla f(\hat{x} + uh) \cdot h = \hat{a}^T \hat{a} + h^T h E(\hat{x}; h)$. 

Since $\nabla f(\tilde{x}) = 0$ at a stationary point, 
$f(\tilde{x} + \tilde{h}) - f(\tilde{x}) = \tilde{h}^T H(\tilde{x}) \tilde{h} + \|\tilde{h}\|^2 \mathbb{E}_2(\tilde{x}; \tilde{h})$.

The main idea is that this term doesn’t matter, and so we are done by the lemma. Here is how this works for (i):

- let $0 < \lambda_1 \leq \ldots \leq \lambda_n$ be eigenvalues of $H(\tilde{x})$.
  - pick $\epsilon > 0$ s.t. $|E_2(\tilde{x}; \tilde{h})| < \frac{1}{4} \lambda$, for $0 < \|\tilde{h}\| < \epsilon$.
  - let $u \in (0, \lambda_1)$, so that the eigenvalues $\lambda_j - u$ of $H(\tilde{x}) - uI_n$ are all positive.
  - Lemma \( \tilde{h}^T [H(\tilde{x}) - uI_n] \tilde{h} > 0 \ \forall \tilde{h} \neq 0 \).
  - So $\tilde{h}^T H(\tilde{x}) \tilde{h} > u \|\tilde{h}\|^2 \ \forall \tilde{h} \neq 0$, and taking $u = \frac{1}{2} \lambda$, gives
    $$\frac{1}{2} \tilde{h}^T H(\tilde{x}) \tilde{h} > \frac{1}{4} \lambda \|\tilde{h}\|^2 \geq \|\tilde{h}\|^2 |E_2(\tilde{x}; \tilde{h})| > 0$$
    \( \Rightarrow f(\tilde{x} + \tilde{h}) - f(\tilde{x}) = \frac{1}{2} \tilde{h}^T H(\tilde{x}) \tilde{h} + \|\tilde{h}\|^2 \mathbb{E}_2(\tilde{x}; \tilde{h}) \geq \frac{1}{2} \tilde{h}^T H(\tilde{x}) \tilde{h} - \|\tilde{h}\|^2 |E_2(\tilde{x}; \tilde{h})| > 0 \ \text{for} \ \tilde{h} \neq 0$, and $f$ has a relative minimum at $\tilde{x}$. \( \square \)

**Lemma on Quadratic Forms:** Let $Q(\tilde{h}) := \tilde{h}^T A \tilde{h}$, $A$ symmetric $n \times n$ with eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. Then

(i) $Q(\tilde{h}) > 0 \ \forall \tilde{h} \neq 0 \ \iff \ \forall j > 0 \ \iff A$ pos.-definite

(ii) $Q(\tilde{h}) < 0 \ \forall \tilde{h} \neq 0 \ \iff \ \forall j < 0 \ \iff A$ neg.-definite

(iii) $Q(\tilde{h})$ takes all values $\iff \lambda_1 < 0 < \lambda_n \ \iff A$ indefinite.

(This is just the Corollary to the Special Theorem, better stated.)

**Theorem 2:** If $f$ is twice continuously differentiable, with a stationary point at $\tilde{x} = \tilde{a}$, then:

(i) $H(\tilde{a})$ pos.-definite $\implies f$ has rel. minimum at $\tilde{a}$

(ii) $H(\tilde{a})$ neg.-definite $\implies f$ has rel. maximum at $\tilde{a}$

(iii) $H(\tilde{a})$ indefinite $\implies f$ has saddle point at $\tilde{a}$.

**Proof:** We record (before erasing) the fact that $f(\tilde{x} + \tilde{h}) - f(\tilde{x}) = \tilde{h}^T \nabla f(\tilde{x}) \tilde{h} + \tilde{h}^T H(\tilde{x}) \tilde{h} + \|\tilde{h}\|^2 \mathbb{E}_2(\tilde{x}; \tilde{h})$. \( \Box \)