Math 51
Lecture 26 – August 14, 2023

Topic(s): determining whether a critical point is a local minimum, local maximum, or a saddle point

Suppose $\mathbf{a}$ is a critical point of $f : \mathbb{R}^n \to \mathbb{R}$, i.e. $\nabla f(\mathbf{a}) = \mathbf{0}$. Then, the quadratic approximation near $\mathbf{a}$ is

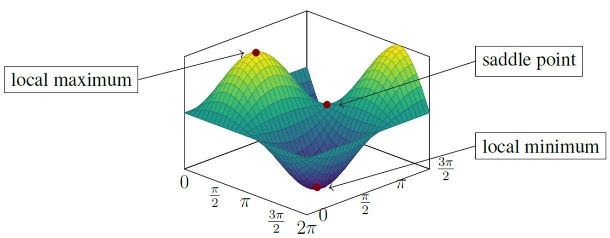
$$f(\mathbf{a}+\mathbf{h}) \approx f(\mathbf{a}) + \frac{1}{2}\mathbf{h}^\intercal Hf(\mathbf{a})\mathbf{h} = f(\mathbf{a}) + \frac{1}{2}q_{Hf(\mathbf{a})}(\mathbf{h}).$$

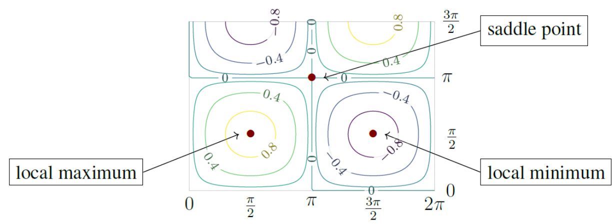
Note that if f is a function from $\mathbb{R}$ to $\mathbb{R}$, then

$$f(a+h) \approx f(a) + \frac{1}{2}f''(a)h^2.$$

Hence, if $f''(a) > 0$, then $f(a+h) > f(a)$ for small values of h, making a a local minimum, and if $f''(a) < 0$, then $f(a+h) < f(a)$ for small values of h, making a a local maximum.

Similarly, if $q_{Hf(\mathbf{a})}$ is positive-definite, then $f(\mathbf{a}+\mathbf{h}) > f(\mathbf{a})$, making $\mathbf{a}$ a local minimum, and if $q_{Hf(\mathbf{a})}$ is negative-definite, then $f(\mathbf{a}+\mathbf{h}) < f(\mathbf{a})$, making $\mathbf{a}$ a local maximum. On the other hand, if $q_{Hf(\mathbf{a})}$ is indefinite, then there are small $\mathbf{h}_1$ and $\mathbf{h}_2$ for which $q_{Hf(\mathbf{a})}(\mathbf{h}_1) > 0$ and $q_{Hf(\mathbf{a})}(\mathbf{h}_2) < 0$. In other words, $f(\mathbf{a}+\mathbf{h}_2) < f(\mathbf{a}) < f(\mathbf{a}+\mathbf{h}_1)$, making $\mathbf{a}$ a saddle point.





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

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

(a) If the Hessian $Hf(\mathbf{a})$ is positive-definite, then $\mathbf{a}$ is a local minimum of f.

(b) If the Hessian $Hf(\mathbf{a})$ is negative-definite, then $\mathbf{a}$ is a local maximum of f.

(c) If the Hessian $Hf(\mathbf{a})$ is indefinite, then $\mathbf{a}$ is a saddle point of f and is neither a local minimum nor a local maximum of f.

If the Hessian $Hf(\mathbf{a})$ is positive-semidefinite or negative-semidefinite (without being definite), then we need more information, just as the single-variable second derivative test is inconclusive when the second derivative vanishes.

It is easy to check the definiteness of diagonal matrices.

Example 1.

$$A = \begin{bmatrix} 11 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \qquad C = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -11 \end{bmatrix}$$

$q_A(x,y,z) = 11x^2 + 3y^2 + 4z^2$ is positive definite.

$q_B(x,y,z) = -3x^2 - y^2 - 3z^2$ is negative definite.

$q_C(x,y,z) = 3x^2 - 4y^2 - 11z^2$ is indefinite.

Example 2. 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?

Since $Hf = \begin{bmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{bmatrix}$, $Hg = \begin{bmatrix} 12x^2 & 0 \\ 0 & -12y^2 \end{bmatrix}$, and $Hh = \begin{bmatrix} -12x^2 & 0 \\ 0 & -12y^2 \end{bmatrix}$,

$$Hf(0,0) = Hg(0,0) = Hh(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which is not definite.

For f, $(0,0)$ is a local minimum since $f(x,y) = x^4 + y^4 \geq 0 = f(0,0)$ for all $(x,y) \in \mathbb{R}^2$.

For g, on $x = 0$, $g(0,y) = -y^4$ and on $y = 0$, $g(x,0) = x^4$. Hence, $(0,0)$ is a saddle point as it is a local minimum on $y = 0$ and a local maximum on $x = 0$.

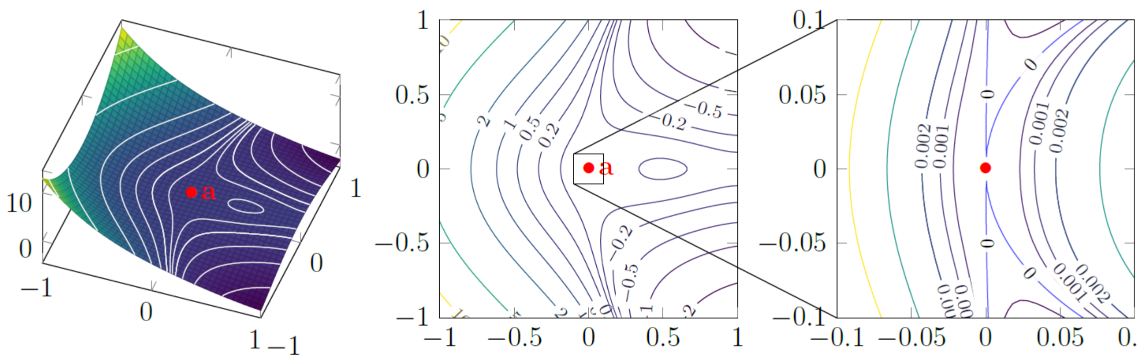
For h, $(0,0)$ is a local maximum since $f(x,y) = -x^4 - y^4 \leq 0 = f(0,0)$ for all $(x,y) \in \mathbb{R}^2$.

Example 3. Consider the function $f(x,y) = (x^2 + y^2 - 3x)^2 + 4x^2(x-2)$. What happens at $(0,0)$?

We compute $\nabla f = \begin{bmatrix} 4x^3 - 6x^2 + 4xy^2 + 2x - 6y \\ 4x^2y - 12xy + 4y^3 \end{bmatrix}$, and so, $(0,0)$ is a critical point.

Next, $Hf = \begin{bmatrix} 12x^2 - 12x + 4y^2 + 2 & 8xy - 12y \\ 8xy - 12y & 4x^2 - 12x + 12y^2 \end{bmatrix}$ and $Hf(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$.

$Hf(0,0)$ has eigenvectors $\vec{w}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{w}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ with eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 0$, respectively. $Hf(0,0)$ is positive semidefinite.



Degenerate point

We can also use the diagonalization formula to determine the nature of a critical point. If $Hf(\mathbf{a})$ has orthogonal eigenvectors $\mathbf{w}_1$ and $\mathbf{w}_2$ (with corresponding eigenvalues λ_1 and λ_2, respectively), then

$$f(\mathbf{a}+h_1\mathbf{w}_1+h_2\mathbf{w}_2) \approx f(\mathbf{a}) + \frac{1}{2}q_{Hf(\mathbf{a})}(h_1\mathbf{w}_1+h_2\mathbf{w}_2) = f(\mathbf{a}) + \frac{1}{2}\lambda_1(\mathbf{w}_1\cdot\mathbf{w}_1)h_1^2 + \frac{1}{2}\lambda_2(\mathbf{w}_2\cdot\mathbf{w}_2)h_2^2.$$

From this, we can tell that the definiteness of the Hessian at $\mathbf{a}$ is determined by the sign of the eigenvalues.

Theorem 26.3.1. (Second Derivative Test, Version 2) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and let $\mathbf{a} \in \mathbb{R}^n$ be a critical point of f.

(a) If the Hessian $Hf(\mathbf{a})$ has all eigenvalues positive, then $\mathbf{a}$ is a local minimum of f.

(b) If the Hessian $Hf(\mathbf{a})$ has all eigenvalues negative, then $\mathbf{a}$ is a local maximum of f.

(c) If the Hessian $Hf(\mathbf{a})$ has both a positive eigenvalue and a negative eigenvalue (some other eigenvalue may be 0), then $\mathbf{a}$ is a saddle point of f.

In the 2×2 case, there is a quick way to determine the signs of the eigenvalues:

Theorem 26.3.3. Consider a symmetric 2×2 matrix $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$, so the eigenvalues are the roots of the characteristic polynomial $P_A(\lambda) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A)$. For its two roots λ_1 and λ_2,

(a) λ_1 and λ_2 have opposite signs precisely when the product $\det(A) = \lambda_1\lambda_2$ is negative; so the indefinite case occurs exactly when $ad - b^2 < 0$.

(b) λ_1 and λ_2 are either both positive or negative precisely when their product $\det(A) = \lambda_1\lambda_2$ is positive.

(c) In case (b), the common sign of λ_1 and λ_2 is the same as that of their sum $\lambda_1 + \lambda_2 = a + d$.

Example 4. Determine the nature of the quadratic form $10x_1^2 - 14x_1x_2 + 5x_2^2$.

The unique symmetric matrix associated is $A = \begin{bmatrix} 10 & -7 \\ -7 & 5 \end{bmatrix}$. Since $\operatorname{tr}(A) = 15$ and $\det(A) = 1$, A is positive-definite.

Example 5. Analyze the behavior of $f(x, y) = x^3 - 3x^2 - 6xy + 9x + 3y^2$ at each of its critical points.

We get $\nabla f = \begin{bmatrix} 3x^2-6x-6y+9 \\ -6x+6y \end{bmatrix}$ and $Hf = \begin{bmatrix} 6x-6 & -6 \\ -6 & 6 \end{bmatrix}$. Hence, the critical points are $(1, 1)$ and $(3, 3)$. At $(1,1)$, $Hf = \begin{bmatrix} 0 & -6 \\ -6 & 6 \end{bmatrix}$ has trace 6 and determinant -36. Hence, $Hf(1,1)$ is indefinite and $(1,1)$ is a saddle point. At $(3,3)$, $Hf = \begin{bmatrix} 12 & -6 \\ -6 & 6 \end{bmatrix}$ has trace 18 and determinant 36. So, $Hf(3,3)$ is positive-definite and $(3,3)$ is a local minimum.

Example 6. Consider the function $f(x,y) = x^3 - 4x^2 + x^2y - 3xy$.

(a) Determine the eigenvalues of the Hessian of f at the critical points, and compute an eigenvector for each eigenvalue for each critical point.

We get $\nabla f = \begin{bmatrix} 3x^2 - 8x + 2xy - 3y \\ x^2 - 3x \end{bmatrix}$ and $Hf = \begin{bmatrix} 6x - 8 + 2y & 2x - 3 \\ 2x - 3 & 0 \end{bmatrix}$. Hence, the critical points are $(0,0)$ and $(3,-1)$.

$Hf(0,0) = \begin{bmatrix} -8 & -3 \\ -3 & 0 \end{bmatrix}$ has eigenvalues $\lambda_1 = -9$ and $\lambda_2 = 1$ with corresponding eigenvectors $\vec{w_1} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\vec{w_2} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$, respectively.

$Hf(3,-1) = \begin{bmatrix} 8 & 3 \\ 3 & 0 \end{bmatrix}$ has eigenvalues $\lambda_1 = -1$ and $\lambda_2 = 9$ with corresponding eigenvectors $\vec{w_1} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ and $\vec{w_2} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, respectively.

(b) Use the information from part (a) to sketch the contour plot of f near each critical point.

$(0,0)$

$\vec{w_1}' = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\vec{w_2}' = \frac{1}{\sqrt{10}}\begin{bmatrix} -1 \\ 3 \end{bmatrix}$.

If $\vec{v} = t_1\vec{w_1}' + t_2\vec{w_2}'$, then

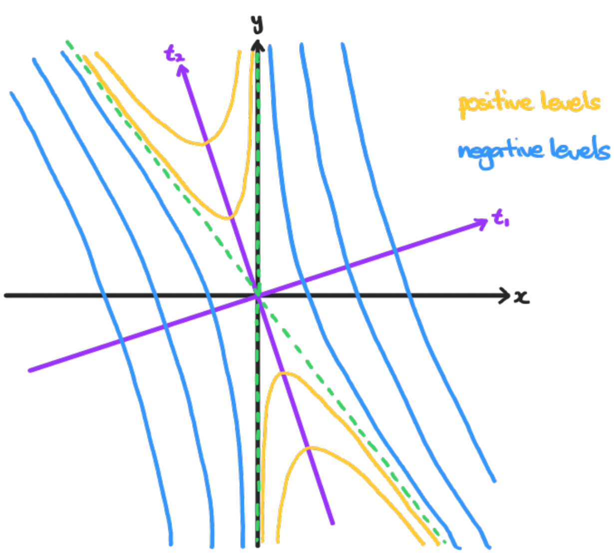
$$q_{Hf(0,0)}(\vec{v}) = -9t_1^2 + t_2^2.$$

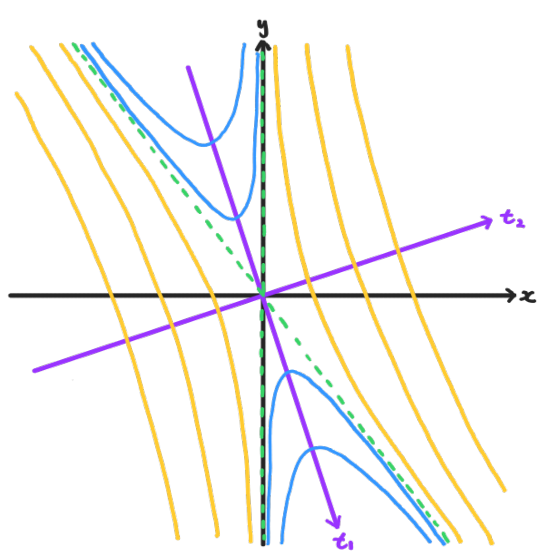
$(3,-1)$

$\vec{w_1}' = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $\vec{w_2}' = \frac{1}{\sqrt{10}}\begin{bmatrix} 3 \\ 1 \end{bmatrix}$

If $\vec{v} = t_1\vec{w_1}' + t_2\vec{w_2}'$, then

$$q_{Hf(3,-1)}(\vec{v}) = -t_1^2 + 9t_2^2.$$





Example 7. Let $f(x, y) = x^3 + 3x^2 + y^2 - 4y$.

(a) Find all critical points of $f(x, y)$.

We compute $\nabla f = \begin{bmatrix} 3x^2+6x \\ 2y-4 \end{bmatrix}$, and so, the critical points are $(0,2)$ and $(-2,2)$.

(b) Compute the Hessian matrix at each of these points, and for each, determine if it is a local maximum, local minimum, or a saddle point.

We compute $Hf = \begin{bmatrix} 6x+6 & 0 \\ 0 & 2 \end{bmatrix}$, $Hf(0,2) = \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix}$ and $Hf(-2,2) = \begin{bmatrix} -6 & 0 \\ 0 & 2 \end{bmatrix}$.

Hence, $(0,2)$ is a local minimum, and $(-2,2)$ is a saddle point.

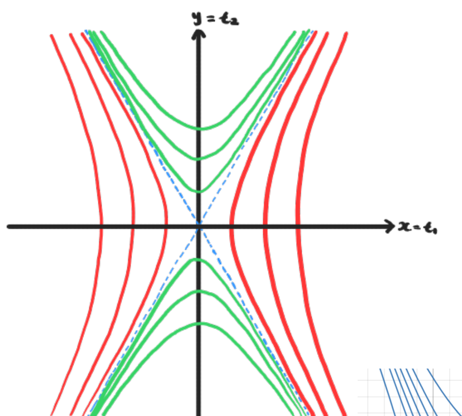
(c) Sketch the contour plot of f near each critical point.

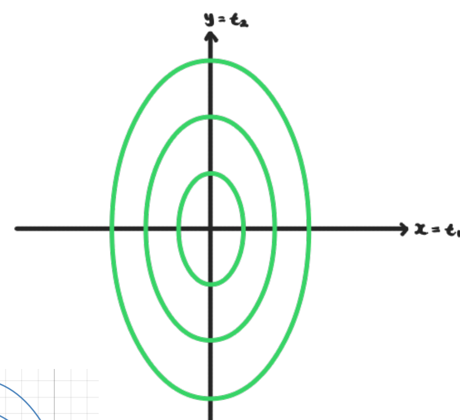
$Hf(-2,2)$ has eigenvectors $\vec{w}_1' = \vec{e}_1$ and $\vec{w}_2' = \vec{e}_2$ with respective eigenvalues $\lambda_1 = -6$ and $\lambda_2 = 2$. For $\vec{v} = t_1\vec{w}_1' + t_2\vec{w}_2'$,

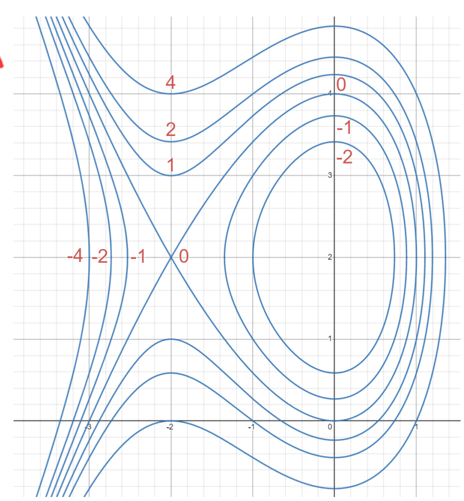
$$q_{Hf(-2,2)}(\vec{v}) = -6t_1^2 + 2t_2^2.$$

$Hf(0,2)$ has eigenvectors $\vec{w}_1' = \vec{e}_1$ and $\vec{w}_2' = \vec{e}_2$ with respective eigenvalues $\lambda_1 = 6$ and $\lambda_2 = 2$. For $\vec{v} = t_1\vec{w}_1' + t_2\vec{w}_2'$,

$$q_{Hf(0,2)}(\vec{v}) = 6t_1^2 + 2t_2^2.$$







Example 8. Find the critical points of $g(x,y) = 3x^2y + 2y^3 - xy$ and classify each as a local maximum, local minimum, or a saddle point.

We compute $\nabla g = \begin{bmatrix} 6xy - y \\ 3x^2 + 6y^2 - x \end{bmatrix}$ and $Hg = \begin{bmatrix} 6y & 6x-1 \\ 6x-1 & 12y \end{bmatrix}$. Setting $\nabla g = \vec{0}$ gives us critical points $P(0,0)$, $Q(\frac{1}{3}, 0)$, $R(\frac{1}{6}, \frac{1}{6\sqrt{2}})$, $S(\frac{1}{6}, -\frac{1}{6\sqrt{2}})$.

$Hg(P) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ has eigenvalues 1 and -1. P is a saddle point.

$Hg(Q) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has eigenvalues 1 and -1. Q is a saddle point.

$Hg(R) = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$ has eigenvalues $\frac{1}{\sqrt{2}}$ and $\sqrt{2}$. R is a local minimum.

$Hg(S) = \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 \\ 0 & -\sqrt{2} \end{bmatrix}$ has eigenvalues $-\frac{1}{\sqrt{2}}$ and $-\sqrt{2}$. S is a local maximum.