Math 5BI: Problem Set 6
Gradient dynamical systems

Recall that if $f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$ is a smooth function of n variables, the *gradient* of f is the vector field

$$\nabla f(\mathbf{x}) = (\nabla f)(x_1, x_2, \dots, x_n) = \begin{pmatrix} (\partial f/\partial x_1)(x_1, x_2, \dots, x_n) \\ \cdots \\ (\partial f/\partial x_n)(x_1, x_2, \dots, x_n) \end{pmatrix},$$

a vector field which is perpendicular to the level sets of f. We say that a point $\mathbf{c} = (c_1, \dots, c_n)$ is a **critical point** for f if $\nabla f(\mathbf{c}) = 0$. Critical points are candidates for maxima and minima.

Problem 6.1. a. Find the critical points of the function $f(x, y) = 3x^2 - 3y^2 - 2x^3$.

b. Find the critical points of the function $f(x, y) = (1/2)y^2 - \cos x$.

We want to investigate the behaviour of a function $f(x_1, \dots, x_n)$ near a critical point $\mathbf{c} = (c_1, \dots, c_n)$ and develop a "second derivative test" for local minima and maxima. To do this, we consider the **Hessian matrix** of all second-order partial derivatives at $\mathbf{c}$:

$$A = \begin{pmatrix} \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_1}\right)(\mathbf{c}) & \cdots & \frac{\partial}{\partial x_n}\left(\frac{\partial f}{\partial x_1}\right)(\mathbf{c}) \\ \cdots & \cdots & \cdots \\ \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_n}\right)(\mathbf{c}) & \cdots & \frac{\partial}{\partial x_n}\left(\frac{\partial f}{\partial x_n}\right)(\mathbf{c}) \end{pmatrix}$$

Now it is a theorem that

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right).$$

Hence the Hessian matrix is always symmetric, $A = A^T$.

Problem 6.2. a. Calculate the Hessian matrix of the function $f(x, y) = 3x^2 - 3y^2 - 2x^3$ at the critical point $(1, 0)$.

b. Calculate the Hessian matrix of the function $f(x, y) = 3x^2 - 3y^2 - 2x^3$ at the critical point $(0, 0)$.

c. Calculate the Hessian matrix of the function $f(x, y) = (1/2)y^2 - \cos x$ at the critical point $(0, 0)$.

d. Calculate the Hessian matrix of the function $f(x,y) = (1/2)y^2 - \cos x$ at the critical point $(\pi, 0)$.

Recall that the eigenvalues of a square matrix A are the solutions λ to the equation

$$\det(A - \lambda I) = 0. \tag{1}$$

Problem 6.3. a. Show that the eigenvalues of a 2×2 *symmetric* matrix with real entries are real.

b. If a 2×2 *symmetric* matrix has two distinct eigenvalues λ_1 and λ_2 show that the corresponding eigenspaces

$$W_{\lambda_1} = \{\mathbf{x} \in \mathbb{R}^2 : A\mathbf{x} = \lambda_1 \mathbf{x}\}, \quad W_{\lambda_2} = \{\mathbf{x} \in \mathbb{R}^2 : A\mathbf{x} = \lambda_2 \mathbf{x}\}$$

are perpendicular to each other. Hint: Use the fact that $\mathbf{x} \cdot \mathbf{y} = \mathbf{y}^t \mathbf{x}$ where on the right side of this equality we are using matrix multiplication. Also use the fact that if A is symmetric, $A^t = A$.

More generally, if A is an $n \times n$ symmetric matrix, it can be proven that all of its eigenvalues are real and that eigenspaces for distinct eigenvalues are perpendicular. In fact, it can be shown that there is a matrix B such that $B^T = B$ and

$$B^T A B = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues of A.

Definition. The symmetric matrix A is said to be

- *positive definite* if all of its eigenvalues are positive.
- *negative definite* if all of its eigenvalues are negative.
- *nondegenerate* if all of its eigenvalues are nonzero.
- *nondegenerate of index* k if it is nondegenerate and exactly k of its eigenvalues are negative.

The second derivative test. *Suppose that $f(x_1, \ldots, x_n)$ has continuous second partial derivatives and $\mathbf{c}$ is a critical point for f. If the Hessian of f at $\mathbf{c}$ is*

1. *positive-definite, then $\mathbf{c}$ is a local minimum,*
2. *negative-definite, then $\mathbf{c}$ is a local maximum,*

If the Hessian of f at $\mathbf{c}$ is nondegenerate of index k, we say that $\mathbf{c}$ is a "saddle point" of index k.

Problem 6.4. a. Which of the critical points of the function $f(x, y) = 3x^2 - 3y^2 - 2x^3$ are local minima? local maxima? saddle points of index one?

b. Which of the critical points of the function $f(x, y) = (1/2)y^2 - \cos x$ are local minima? local maxima? saddle points of index one?

c. Which of the critical points of the function $f(x, y) = \cos x - (1/2)y^2$ are local minima? local maxima? saddle points of index one?

How do we see that the second derivative test works? If $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$, we can regard the gradient of f as defining a system of differential equations

$$\begin{array}{rcl} \frac{dx_1}{dt} & = & \frac{\partial f}{\partial x_1}(x, x_2, \ldots, x_n) \\ \cdots & & \cdots \\ \frac{dx_n}{dt} & = & \frac{\partial f}{\partial x_n}(x, x_2, \ldots, x_n) \end{array} \qquad (2)$$

Such a system of differential equations is called a *gradient dynamical system*. It can be written in vector form as

$$\frac{d\mathbf{x}}{dt} = \nabla f(\mathbf{x}).$$

A constant solution $\mathbf{c} = (c_1, \ldots, c_n)$ to the gradient dynamical system (2) is just a critical point for f. It is easy to visualize gradient dynamical systems in two variables. One begins by plotting the level curves $f(x_1, x_2) = c$, thus obtaining a *topographic map* of the surface $z = f(x_1, x_2)$. The orbits of the gradient dynamical system are then just the orbits of the gradient dynamical system.

One can think of the orbits of the gradient dynamical system

$$\frac{d\mathbf{x}}{dt} = \nabla f(\mathbf{x})$$

as representing the paths of rain droplets flowing over the surface $z = f(x_1, x_2)$, except that they are traversed in the opposite direction. The mountain peaks, mountain passes, and lake bottoms on the topographic map are included among the critical points of f.

In more than two variables, the orbits of such systems are still orthogonal to the level sets $f(x_1, \ldots, x_n) = c$. One can have the same geometrical picture in one's mind.

To investigate the behaviour of a function $f(x_1, \ldots, x_n)$ near a critical point $\mathbf{c} = (c_1, \ldots, c_n)$, we can consider the **linearization** of the gradient dynamical system (2) at the equilibrium solution $\mathbf{c}$:

$$\begin{pmatrix} dx_1/dt \\ \cdots \\ dx_n/dt \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_1}\right)(\mathbf{c}) & \cdots & \frac{\partial}{\partial x_n}\left(\frac{\partial f}{\partial x_1}\right)(\mathbf{c}) \\ \cdots & \cdots & \cdots \\ \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_n}\right)(\mathbf{c}) & \cdots & \frac{\partial}{\partial x_n}\left(\frac{\partial f}{\partial x_n}\right)(\mathbf{c}) \end{pmatrix} \begin{pmatrix} x_1 - c_1 \\ \cdots \\ x_n - c_n \end{pmatrix}.$$

If we let

$$a_{ij} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\mathbf{c}),$$

we can rewrite this system as

$$\begin{pmatrix} dx_1/dt \\ \cdots \\ dx_n/dt \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 - c_1 \\ \cdots \\ x_n - c_n \end{pmatrix},$$

or equivalently, as

$$\frac{d\mathbf{x}}{dt} = A(\mathbf{x} - \mathbf{c}) \quad \text{or} \quad \frac{d\mathbf{y}}{dt} = A\mathbf{y}, \quad \text{where} \quad \mathbf{y} = \mathbf{x} - \mathbf{c},$$

and

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

is the Hessian matrix.

If $\mathbf{c}$ is a critical point for $f(x_1, \dots, x_n)$ and

$$\frac{d\mathbf{x}}{dt} = A(\mathbf{x} - \mathbf{c}) \quad \text{is the linearization of} \quad \frac{d\mathbf{x}}{dt} = \nabla f(\mathbf{x})$$

at $\mathbf{c}$, the eigenvalues of A determine the qualitative behaviour of the solutions to the linearization. If all of the eigenvalues of A are negative, then all the nonzero solutions will tend towards $\mathbf{c}$ as $t \to \infty$. We see that in this case $\mathbf{c}$ is a **local maximum**. If all of the eigenvalues are positive, then all the nonzero solutions will move away from $\mathbf{c}$ as $t \to \infty$ and $\mathbf{c}$ musts be a **local minimum**.