MATH 162: Calculus II
Framework for Thurs., Mar. 29–Fri. Mar. 30
The Gradient Vector

Today's Goal: To learn about the gradient vector $\vec{\nabla} f$ and its uses, where f is a function of two or three variables.

The Gradient Vector

Suppose f is a differentiable function of two variables x and y with domain R, an open region of the xy-plane. Suppose also that

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}, \qquad t \in I,$$

(where I is some interval) is a differentiable vector function (parametrized curve) with $(x(t), y(t))$ being a point in R for each $t \in I$. Then by the chain rule,

$$\begin{aligned} \frac{d}{dt} f(x(t), y(t)) &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= [f_x\mathbf{i} + f_y\mathbf{j}] \cdot [x'(t)\mathbf{i} + y'(t)\mathbf{j}] \\ &= [f_x\mathbf{i} + f_y\mathbf{j}] \cdot \frac{d\mathbf{r}}{dt}. \end{aligned} \tag{1}$$

Definition: For a differentiable function $f(x_1, \ldots, x_n)$ of n variables, we define the *gradient vector of f* to be

$$\vec{\nabla} f \;:=\; \left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right\rangle.$$

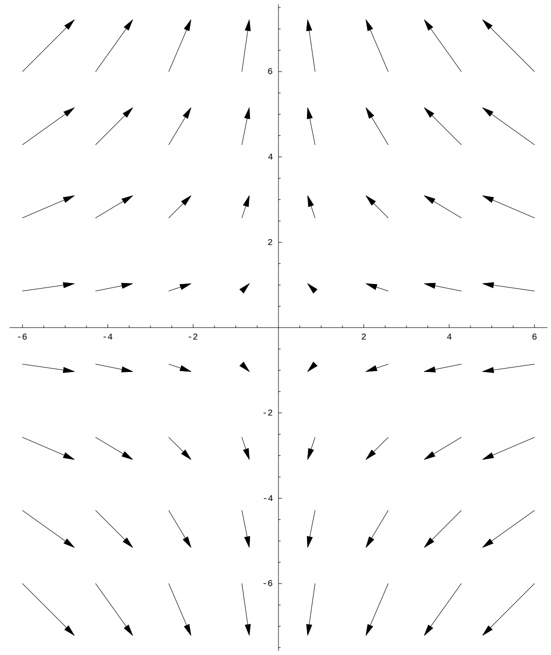
Remarks:

- Using this definition, the total derivative df/dt calculated in (1) above may be written as
$$\frac{df}{dt} \;=\; \vec{\nabla} f \cdot \mathbf{r}'.$$
In particular, if $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$, $t \in (a, b)$ is a differentiable vector function, and if f is a function of 3 variables which is differentiable at the point (x_0, y_0, z_0), where $x_0 = x(t_0)$, $y_0 = y(t_0)$, and $z_0 = z(t_0)$ for some $t_0 \in (a, b)$, then
$$\left.\frac{df}{dt}\right|_{t=t_0} \;=\; \vec{\nabla} f(x_0, y_0, z_0) \cdot \mathbf{r}'(t_0).$$

- If f is a function of 2 variables, then $\vec{\nabla} f$ has 2 components. Thus, while the graph of such an f lives in 3D, $\vec{\nabla} f$ should be thought of as a vector in the plane.

  If f is a function of 3 variables, then $\vec{\nabla} f$ has 3 components, and is a vector in 3-space.

Speaking more generally, we may say that while a function $f(x_1, \ldots, x_n)$ of n variables requires n inputs to produce a single (numeric) output, the corresponding gradient $\vec{\nabla} f$ produces from those same n inputs a *vector* with n components. Objects which assign to each n-tuple input an n-vector output are known as *vector fields*. The gradient is an example of a vector field.



Example: For $f(x, y) = y^2 - x^2$, we have $\text{dom}(f) = \mathbb{R}^2$ and $\vec{\nabla} f(x, y) = -2x\mathbf{i} + 2y\mathbf{j}$. Selecting any point (x, y) in the plane, we may choose to draw $\vec{\nabla} f(x, y)$ not as a vector in standard position, but rather one with initial point (x, y), obtaining the picture at right.

- **Properties of the gradient operator**: If f, g are both differentiable functions of n variables on an open region R, and c is any real number (constant), then

  1. $\vec{\nabla}(cf) = c\vec{\nabla} f$ (constant multiple rule)

  2. $\vec{\nabla}(f \pm g) = \vec{\nabla} f \pm \vec{\nabla} g$ (sum/difference rules)

  3. $\vec{\nabla}(fg) = g(\vec{\nabla} f) + f(\vec{\nabla} g)$ (product rule)

  4. $\vec{\nabla}\left(\dfrac{f}{g}\right) = \dfrac{g(\vec{\nabla} f) - f(\vec{\nabla} g)}{g^2}$ (quotient rule)

Directional Derivatives

Let $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ be a unit vector (i.e., $|\mathbf{u}| = 1$), and let $\mathbf{r}(t)$ be the vector function

$$\mathbf{r}(s) = (x_0 + su_1)\mathbf{i} + (y_0 + su_2)\mathbf{j},$$

which parametrizes the line through $P = (x_0, y_0)$ parallel to $\mathbf{u}$ in such a way that the "speed" $|d\mathbf{r}/dt| = |\mathbf{u}| = 1$. We make the following definition.

Definition: For a function f of two variables that is differentiable at (x_0, y_0), we define the *directional derivative of f at (x_0, y_0) in the direction* $\mathbf{u}$ to be

$$D_{\mathbf{u}} f(x_0, y_0) \ := \ \lim_{x \to 0} \frac{f(x_0 + su_1, y_0 + su_2) - f(x_0, y_0)}{s} \ = \ \left(\frac{df}{ds}\right)_{\mathbf{u},P} = \vec{\nabla} f(x_0, y_0) \cdot \mathbf{u}.$$

Remarks:

- **Special cases**: the partial derivatives f_x, f_y themselves are derivatives in the directions $\mathbf{i}$, $\mathbf{j}$ respectively.
$$D_{\mathbf{i}}f = \vec{\nabla} f \cdot \mathbf{i} = f_x, \qquad \text{and} \qquad D_{\mathbf{j}}f = \vec{\nabla} f \cdot \mathbf{j} = f_y.$$

- For f a function of 2 variables, the direction of $\vec{\nabla} f(x,y)$ (namely $\vec{\nabla} f(x,y)/|\vec{\nabla} f(x,y)|$) is the direction of maximum increase, while $-\vec{\nabla} f(x,y)/|\vec{\nabla} f(x,y)|$ is the direction of maximum decrease.

  Proof: For any unit vector $\mathbf{u}$,
$$D_{\mathbf{u}}f(x,y) \;=\; \vec{\nabla} f(x,y) \cdot \mathbf{u} \;=\; |\vec{\nabla} f(x,y)||\mathbf{u}|\cos\theta \;=\; |\vec{\nabla} f(x,y)|\cos\theta,$$
  where θ is the angle between $\vec{\nabla} f(x,y)$ and $\mathbf{u}$. This directional derivative is largest when $\theta = 0$ (i.e., when $\mathbf{u} = \vec{\nabla} f/|\vec{\nabla} f|$) and smallest when $\theta = \pi$.

- The notion of directional derivative extends naturally to functions of 3 or more variables.

The Gradient and Level Sets

Suppose $f(x,y)$ is differentiable at the point $P = (x_0, y_0)$, and let $k = f(x_0, y_0)$. Then the k-level curve of f contains (x_0, y_0). Suppose that we have a parametrization of a section of this level curve containing the point (x_0, y_0). That is, let

- $x(t)$ and $y(t)$ be differentiable functions of t in an open interval I containing $t = 0$,
- $x_0 = x(0)$ and $y_0 = x(0)$, and
- $f(x(t), y(t)) = k$ for $t \in I$ (that is, the parametrization gives at least a small part of the k-level curve of f—a part that contains the point P).

Because we are parametrizing a level curve of f, it follows that $df/dt = 0$ for $t \in I$. In particular,
$$0 \;=\; \left.\frac{df}{dt}\right|_{t=t_0} \;=\; \vec{\nabla} f(x_0, y_0) \cdot [x'(t_0)\mathbf{i} + y'(t_0)\mathbf{j}].$$
This shows that the gradient vector at P is orthogonal to the level curve of f (or the tangent line to the level curve) through P. This is true at all points P where f is differentiable. That this result may be generalized to higher dimensions is motivation for the definition of a tangent plane.

Tangent Planes

Definition: Let $f(x, y, z)$ be differentiable at a point $P = (x_0, y_0, z_0)$ contained in the level surface $f(x, y, z) = k$. We define the *tangent plane to this level surface of f at P* to be the plane containing P normal to $\vec{\nabla} f(x_0, y_0, z_0)$.

Example: Suppose $f(x, y, z)$ is a differentiable function at the point $P = (x_0, y_0, z_0)$ lying on the level surface $f(x, y, z) = k$. Derive a formula for the equation of the tangent plane to this level surface of f at P. Then use it to write the equation of the tangent plane to the quadric surface

$$f(x, y, z) = x^2 + 3y^2 + 2z^2 \ = \ 6$$

at the point $(1, 1, 1)$.

Example: Suppose $z = f(x, y)$ is a differentiable function at the point $P = (x_0, y_0)$. Derive a formula for the equation of the tangent plane to the surface $z = f(x, y)$ at the point $P = (x_0, y_0, f(x_0, y_0))$. Use it to get the equation of the tangent plane to $z = 2x^2 - y^2$ at the point $(1, 3, -7)$.