14.6. **Directional derivatives and the gradient vector.**

Definition 14.11. The directional derivative of f at (x_0, y_0) in the direction of a unit vector $u = \langle a, b\rangle$ is

$$D_u f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h}.$$

Remark 14.16. If u has length c and expressed as $u = \langle ac, bc\rangle$, then the limit in the above definition leads to

$$\lim_{h\to 0} \frac{f(x_0 + ach, y_0 + bch) - f(x_0, y_0)}{h} = cD_u(x_0, y_0).$$

Remark 14.17. If $\mathbf{i} = \langle 1, 0\rangle$ and $\mathbf{j} = \langle 0, 1\rangle$, then $f_x(x_0, y_0) = D_{\mathbf{i}} f(x_0, y_0)$ and $f_y(x_0, y_0) = D_{\mathbf{j}} f(x_0, y_0)$.

Example 14.24. Let $f(x, y) = x^2 + y^2$ and $u = \langle 1/2, \sqrt{3}/2\rangle$. Then,

$$D_u f(1, 2) = \lim_{h\to 0} \frac{f(1 + h/2, 2 + \sqrt{3}h/2) - f(1, 2)}{h} = 1 + 2\sqrt{3}.$$

Theorem 14.11. *If $f(x, y)$ is differentiable at (x_0, y_0), then for any unit vector $u = \langle a, b\rangle$, $D_u f(x_0, y_0) = af_x(x_0, y_0) + bf_y(x_0, y_0)$.*

Proof. Since f is differentiable at (x_0, y_0),

$$f(x_0 + ah, y_0 + bh) = f(x_0, y_0) + f_x(x_0, y_0)ah + f_y(x_0, y_0)bh + \epsilon_1 ah + \epsilon_2 bh$$

where $(\epsilon_1, \epsilon_2) \to (0, 0)$ as $h \to 0$. This implies

$$\lim_{h\to 0} \frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h} = af_x(x_0, y_0) + bf_y(x_0, y_0).$$

□

Example 14.25. Let $f(x, y) = e^{x \sin \pi y}$ and $u = \langle 1/2, \sqrt{3}/2\rangle$. Then, $f_x = e^{x \sin \pi y} \sin \pi y$ and $f_y = e^{x \sin \pi y} \pi x \cos \pi y$. Since f_x and f_y is continuous on $\mathbb{R}^2$, f is differentiable on $\mathbb{R}^2$ and

$$D_u f(1, 0) = \frac{1}{2} f_x(1, 0) + \frac{\sqrt{3}}{2} f_y(1, 0) = \frac{\sqrt{3}\pi}{2}.$$

Example 14.26. Let $f(x, y) = xy^2/(x^2 + y^2)$ for $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$. It has been proved before that f is continuous at $(0, 0)$. For unit vector $u = \langle a, b\rangle$, one may compute

$$D_u f(0, 0) = \lim_{h\to 0} \frac{f(ah, bh) - f(0, 0)}{h} = ab^2.$$

This implies that $D_u f(0, 0) \neq 0$ for $a \neq 0$ and $b \neq 0$. Note that $f_x(0, 0) = f_y(0, 0) = 0$ and this yields $af_x(0, 0) + bf_y(0, 0) = 0$. As $D_u f(0, 0) \neq af_x(0, 0) + bf_y(0, 0)$, f is not differentiable at $(0, 0)$. It is remarkable that the existence of directional derivatives in all directions is not sufficient for the differentiability.

Definition 14.12. The gradient of $f(x, y)$ is the vector function $\nabla f(x, y)$ or $\mathbf{grad} f(x, y)$ defined by $\nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle$.

Remark 14.18. If f is a differentiable function at (x_0, y_0), then $D_u f(x_0, y_0) = \nabla f(x_0, y_0) \cdot u$.

Example 14.27. For function $f(x, y) = e^{x \sin \pi y}$, $\nabla f(x, y) = f(x, y)\langle \sin \pi y, \pi x \cos \pi y\rangle$.

Remark 14.19. For functions of n variables say $f(x_1, ..., x_n)$, the directional derivative of f at $\mathbf{x} = (x_1, ..., x_n)$ in the direction of a unit vector $u = \langle a_1, ..., a_n \rangle$ is defined by

$$D_u f(\mathbf{x}) = \lim_{h \to 0} \frac{f(x_1 + a_1 h, ..., x_n + a_n h) - f(x_1, ..., x_n)}{h}.$$

The gradient of f at $\mathbf{x}$ is defined to be the vector $\nabla f(\mathbf{x}) = \langle f_{x_1}(\mathbf{x}), ..., f_{x_n}(\mathbf{x}) \rangle$. When f is differentiable at $\mathbf{x}$, then $D_u f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot u$.

Example 14.28. For $f(x, y, z) = \sin xyz$, $\nabla f = \langle yz \cos xyz, xz \cos xyz, xy \cos xyz \rangle$ and the directional derivative of f at $(1, 1, \pi)$ in the direction $u = \langle 1/2, 1/2, 1/\sqrt{2} \rangle$ is

$$\nabla f(1, 1, \pi) \cdot u = \langle -\pi, -\pi, -1 \rangle \cdot \langle 1/2, 1/2, 1/\sqrt{2} \rangle = -\pi - 1/\sqrt{2}.$$

Theorem 14.12. *Let $f(\mathbf{x})$ be differentiable with $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n$ and $U = \{u \in \mathbb{R}^n : |u| = 1\}$. Then,* $\max\{|D_u f(\mathbf{x})| : u \in U\} = |D_v f(\mathbf{x})| = |D_{-v} f(\mathbf{x})| = |\nabla f(\mathbf{x})|$, *where* $v = \nabla f(\mathbf{x}) / |\nabla f(\mathbf{x})|$.

Proof. Note that $D_u f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot u = |\nabla f(\mathbf{x})| \cos \theta$, where θ is the angle between $\nabla f(\mathbf{x})$ and u. This implies that $|D_u f(\mathbf{x})|$ achieves its maximum when $\theta = 0$ or π, that is, $u = v$ or $u = -v$, where $v = \nabla f(\mathbf{x}) / |\nabla f(\mathbf{x})|$. In this case, we have $|D_v f(\mathbf{x})| = |\nabla f(\mathbf{x})|$. □

Example 14.29. Let $f(x, y) = xe^y$ and $P = (1, 0)$. Then, the maximum rate of change of f at P equals $|\nabla f(P)| = |\langle 1, 1 \rangle| = \sqrt{2}$ in the direction of $\nabla f(P) = \langle 1, 1 \rangle$.

The gradient vector is closely related to the level curve. Consider the level curve $f(x, y) = k$, where k is a constant. Assume that the curve has parametrization $r(t) = \langle x(t), y(t) \rangle$. This implies $f(x(t), y(t)) = k$ and then

$$0 = f_x(r(t)) x'(t) + f_y(r(t)) y'(t) = \nabla f(r(t)) \cdot r'(t).$$

It is easy to see from the above equality that the gradient vector and the tangent line of a level curve are perpendicular.

For functions of three variables, let $r(t) = \langle x(t), y(t), z(t) \rangle$ be a parametric curve on the level surface $F(x, y, z) = k$, where k is a constant. As before, we have $F(r(t)) = k$ and $\nabla F(r(t)) \cdot r'(t) = 0$. Consequently, when $F(x_0, y_0, z_0) = k$ and the level surface $F(x, y, z) = k$ has a tangent plane at (x_0, y_0, z_0), $\nabla F(x_0, y_0, z_0)$ is the normal vector to the tangent plane and thus the tangent plane is given by $\nabla F(x_0, y_0, z_0) \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0$ or equivalently

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0.$$

Immediately, the normal line to the surface $F(x, y, z) = k$ at (x_0, y_0, z_0) is $\{(x_0, y_0, z_0) + t \nabla F(x_0, y_0, z_0) | t \in \mathbb{R}\}$. Particularly, if $F(x, y, z) = f(x, y) - z$ and $k = 0$, then the tangent plane to $F(x, y, z) = 0$ (i.e. $z = f(x, y)$) at (x_0, y_0, z_0) with $z_0 = f(x_0, y_0)$ is given by

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0,$$

which is exactly the same as before.

Example 14.30. Let $F(x, y, z) = x^2 - y^2 + z^2$. Note that $\nabla F(x, y, z) = 2\langle x, -y, z \rangle$ and $F(1, 1, 0) = 0$. Then, the tangent plane to the surface $F(x, y, z) = 0$ at $(1, 1, 0)$ has equation $x = y$ and the normal line is parametrized by $x = 1 + t$, $y = 1 - t$ and $z = 0$ for $t \in \mathbb{R}$.