LESS MUNDANE APPLICATIONS OF THE MOST MUNDANE FUNCTIONS

PISHENG DING

Abstract. Linear functions are arguably the most mundane among all functions. However, the basic fact that a multi-variable linear function has a constant gradient field can provide simple geometric insights into several familiar results such as the Cauchy-Schwarz inequality, the GM-AM inequality, and some distance formulae, as we shall show.

1. Preliminaries

We set the stage for the applications by recalling some basics. For \( f(x, y) = ax + by \), \( \nabla f \equiv a\mathbf{i} + b\mathbf{j} \). As a result, at any point in \( \mathbb{R}^2 \), it is in the direction \( a\mathbf{i} + b\mathbf{j} \) that the directional derivative attains its maximum value of \( |\nabla f| = \sqrt{a^2 + b^2} \). Geometrically, this means that the slope of the planar graph of \( f \) equals \( |\nabla f| = \sqrt{a^2 + b^2} \). The level sets of \( f \) are the parallel lines perpendicular to \( a\mathbf{i} + b\mathbf{j} \). Denote the level-0 line of \( f \) by \( L_0 \). Any point in \( \mathbb{R}^2 \) can be reached by traveling from a point on \( L_0 \) in either the normal direction \( a\mathbf{i} + b\mathbf{j} \) or its opposite for a certain distance. We can thus assign, to any point \( P \in \mathbb{R}^2 \), its signed distance \( \delta(P) \) to \( L_0 \); \( \delta(P) > 0 \) iff \( P \) is in the open half-plane bordered by \( L_0 \) into which \( \nabla f \) points. A key observation is that

\[
(1.1) \quad f(P) = |\nabla f| \cdot \delta(P).
\]

This is the principle “rise = slope \times run” – the essence of linearity.

An entirely analogous situation takes place in the case of a three-variable linear function \( F(x, y, z) = ax + by + cz \), for which the level sets are the planes normal to \( \nabla F \equiv a\mathbf{i} + b\mathbf{j} + c\mathbf{k} \). For a point \( P \in \mathbb{R}^3 \), we similarly define its signed distance \( \delta(P) \) to \( F \)’s level-0 plane \( \Pi_0 \). Then,

\[
(1.2) \quad F(P) = |\nabla F| \cdot \delta(P).
\]

Equations (1.1) and (1.2) hold the key to the ensuing applications.

2. Calculating distances without any formulae

What is the distance between the two lines \( L_1 := \{(x, y) \mid 3x + 4y = 1\} \) and \( L_{31} := \{(x, y) \mid 3x + 4y = 31\} \)? Even if you remember a relevant formula, you might prefer this one-liner: \( L_1 \) and \( L_{31} \) are the level-1 and level-31 sets of \( f(x, y) := 3x + 4y \) and therefore the distance (the “run”) between them equals

\[
\frac{\text{the rise}}{\text{the slope}} = \frac{31 - 1}{|\nabla f|} = \frac{30}{5} = 6.
\]

2020 Mathematics Subject Classification. 26D99, 26B25, 52A40.

Keywords and phrases. gradient, Cauchy-Schwarz inequality, GM-AM inequality, convex functions, distance formula.
Similarly, the distance between the two planes $\Pi_1 := \{ (x, y, z) \mid 2x + y + 2z = 1 \}$ and $\Pi_{31} := \{ (x, y, z) \mid 2x + y + 2z = 31 \}$ equals
\[
\frac{31 - 1}{|\nabla(2x + y + 2z)|} = \frac{30}{3} = 10.
\]

What is the distance between a point and a plane, for example the point $P_0 := (1, 2, 1)$ and the plane $\Pi_0 := \{ (x, y, z) \mid 2x + 3y + 6z = 0 \}$? Note that $\Pi_0$ is the level-0 plane of $F(x, y, z) := 2x + 3y + 6z$ and the distance sought therefore equals
\[
\frac{\text{the change}}{\text{the rate}} = \frac{|F(P_0) - F(\Pi_0)|}{|\nabla F|} = \frac{14 - 0}{7} = 2.
\]

Indeed, the relevant distance formulae can be derived by the same arguments.

3. **Deducing the Cauchy-Schwartz inequality by inspection**

For two $n$-tuples of real numbers $(a_i)_{i=1}^n$ and $(x_i)_{i=1}^n$, the Cauchy-Schwartz inequality asserts that $\sum_{i=1}^n a_i x_i \leq \sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n x_i^2}$. For its proof, it suffices to treat the special case in which $\sum_{i=1}^n x_i^2 = 1$; for, if $\sum_{i=1}^n x_i^2 = s^2$, then $\sum_{i=1}^n (x_i/s)^2 = 1$.

We motivate our proof through an example. Let $S^2 = \{ (x, y, z) \mid x^2 + y^2 + z^2 = 1 \}$, the unit 2-sphere; let $F(x, y, z) = 2x + 3y + 6z$. Consider this familiar problem: Find $\max_{S^2} F$ and $\min_{S^2} F$. In light of Eq. (1.2), the solution can be argued by mere inspection, almost without words, as follows.

View the level-0 plane $\Pi_0$ of $F$ as $S^2$’s equatorial plane and its normal direction $2i + j + 2k$ as North. By Eq. (1.2), the extrema of $F$ on $S^2$ are attained precisely at the points on $S^2$ that are farthest from $\Pi_0$, i.e., the North Pole $P_+$ and the South Pole $P_-$. Clearly, $P_{\pm} = (\pm \frac{2}{\sqrt{7}}, \pm \frac{1}{\sqrt{7}}, \pm \frac{2}{\sqrt{7}})$, at which Eq. (1.2) again gives the $F$-value: $F(P_{\pm}) = |\nabla F| \cdot \delta(P_{\pm}) = 3 \cdot (\pm 1)$. Hence, $|2x + y + 2z| \leq 3$ for $(x, y, z) \in S^2$.

This argument applied to the linear function $ax + by + cz$ yields that
\[
|ax + by + cz| \leq |\nabla(ax + by + cz)| = \sqrt{a^2 + b^2 + c^2} \quad \text{for} \quad (x, y, z) \in S^2,
\]
which, as remarked earlier, implies the Cauchy-Schwartz inequality for $n = 3$.

![Figure 1. A hyperplane $H$ and the unit $(n-1)$-sphere in $\mathbb{R}^n$.](image)

The preceding argument even works in higher dimensions. Let $G$ be a homogeneous degree-1 polynomial in $x_1, \cdots, x_n$. Its level-0 set $H$ is a hyperplane (a 1-codimensional subspace) in $\mathbb{R}^n$. By an orthogonal transformation of coordinates, we may install a new coordinate system $(x'_1, \cdots, x'_n)$ such that $H$ is the subspace
$x'_n = 0$ and that, for $P \in \mathbb{R}^n$, its primed $n$th coordinate $x'_n(P)$ is its signed distance to $H$. See Figure 1. We thus have

$$G(P) = |\nabla G| \cdot x'_n(P).$$

The unit $(n - 1)$-sphere $S^{n-1}$ in $\mathbb{R}^n$ satisfies the same equation in the primed coordinates: $\sum_{i=1}^{n} (x'_i)^2 = 1$, from which it is clear that $x'_n$ on $S^{n-1}$ attains its extrema at precisely the two points $P_{\pm}$ whose primed coordinates are $(0, \cdots, 0, \pm 1)$, implying that $|G(P)| \leq |\nabla G|$ for $P \in S^{n-1}$.

4. Visualizing the GM-AM inequality

For $x_1, \cdots, x_n > 0$, the GM-AM inequality asserts that their geometric mean $\sqrt[n]{x_1 \cdots x_n}$ is no greater than their arithmetic mean $(x_1 + \cdots + x_n)/n$. The following special case is an equivalent statement:

$$x_1 + \cdots + x_n \geq n$$

for $x_1, \cdots, x_n > 0$ such that $\prod_{i=1}^{n} x_i = 1$.

For, if $\prod_{i=1}^{n} x_i = p$, then $\prod_{i=1}^{n} \left( x_i/\sqrt[n]{p} \right) = 1$. We show a visual argument for this statement in the case $n = 2$ and suggest one for $n = 3$.

![Figure 2](image1.png)

**Figure 2.** The line $x + y = 0$; the curve $xy = 1$ with $x, y > 0$.

![Figure 3](image2.png)

**Figure 3.** The plane $x + y + z = 0$; the surface $xyz = 1$ with $x, y, z > 0$.

Let $C = \{(x, y) \mid x, y > 0; \ xy = 1\}$. Where on $C$ does $x + y$ attain its least value? The answer is given by Eq. (1.1): at those points on $C$ that are closest to the
line $L_0 = \{(x, y) \mid x + y = 0\}$. It is apparent from Figure 2 that $(1, 1)$ is the only minimizer of this distance and therefore $x + y \geq 1 + 1$ for $(x, y) \in C$. For proof, view $C$ as the graph of a convex function $h$ defined on $L_0$. (In a new coordinate system making a $45^\circ$-angle with the standard one, an equation for $C$ in the new coordinates $(x', y')$ is $y' = \sqrt{1 + (x')^2}$. As $h$ is convex, it attains its global minimum exactly once; see [1] for convex functions. Symmetry then dictates where the minimizer is.

For the case $n = 3$, let $S = \{(x, y, z) \mid x, y, z > 0; \ xyz = 1\}$, which is the graph of the function $g(x, y) = (xy)^{-1}$ on the first quadrant $(0, \infty)^2 \subset \mathbb{R}^2$. On $S$, the quantity $x + y + z$ attains its least value at a point closest to the plane $\Pi_0 = \{(x, y, z) \mid x + y + z = 0\}$ (by (1.2)). Looking at Figure 3, it is plausible that $(1, 1, 1)$ is the unique minimizer of this distance, which then implies that $x + y + z \geq 1 + 1 + 1$ for $(x, y, z) \in S$. In fact, this situation is almost identical to the previous case: $S$ can be viewed as the graph of a convex function on $\Pi_0$; convexity then entails the existence and uniqueness of a minimizer and symmetry dictates its location. To make this visual argument rigorous, we must analytically check that $S$ can indeed be viewed as the graph of a function on $\Pi_0$ and that this function is indeed convex. For the first claim, we need to show that $S$ passes the “vertical line test” where a “vertical line” means a line perpendicular to the domain $\Pi_0$. From a point $(x, y, -x - y) \in \Pi_0$, proceed in the “vertical” direction $\nabla(x + y + z) = e_1 + e_2 + e_3$ to arrive at points of the form

$$P(t) = (x + t, y + t, -x - y + t).$$

For how many values of $t$ will $P(t) \in S$? A computer can check for us that the cubic equation in $t$

$$(x + t)(y + t)(-x - y + t) = 1$$

has exactly one real solution. For the other claim concerning convexity, first note that the function $g(x, y) = (xy)^{-1}$ is convex on the first quadrant; one can verify this by checking that the Hessian quadratic form of $g$ is positive definite at every point in the first quadrant. Hence $S$ lies on one side of each of its tangent planes, a fact that is invariant under a change of perspective.

References

[1] Wendell H. Fleming, *Functions of Several Variables*, Addison-Wesley, Reading, Massachusetts, 1965.