MULTIVARIABLE CALCULUS

OLIVER KNILL, MATH 21A

Lecture 24: The gradient

Gradient theorem

A most important theorem in multi-variable calculus is the **gradient theorem**:

$\nabla f(x_0, y_0)$ is perpendicular to the level curve passing through (x_0, y_0).

The proof is that if the level curve is parametrized as $\vec{r}(t)$, then the function does not change so that $0 = \frac{d}{dt} f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$ which means that the gradient $\nabla f(\vec{r}(t))$ is perpendicular to the velocity $\vec{r}\,'(t)$. In higher dimensions, the proof is similar, just check it for every curve.

Angles dance upwards

A most important paradigm in **machine learning** is that we can improve a function by going into the direction of the gradient. This is the **angel theorem** because angels dance upwards:

The function f increases most in the gradient direction.

The proof also follows from the chain rule. We know

$$\frac{d}{dt} f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t) = |\nabla f(\vec{r}(t))| \vec{r}\,'(t)| \cos(\alpha) \; .$$

This is biggest if $\alpha = 0$. This means if $\vec{r}\,'(t)$ points into the same direction than ∇f.

Quadratic approximation

We have seen the linear approximation $L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$. The next best improvement is the quadratic approximation. In one dimensions it is

$$Q(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)(x - x_0)^2}{2} \; .$$

In two dimensions, all the 4 derivatives $f_{xx}, f_{xy}, f_{yx}, f_{yy}$ appear. Because of Clairaut $f_{xy} = f_{yx}$, we have

$$Q(x, y) = L(x, y) + \frac{[f_{xx}(x_0, y_0)(x - x_0)^2 + 2 f_{xy}(x_0, y_0)(x - x_0)(y - y_0) + f_{yy}(x_0, y_0)(y - y_0)^2]}{2} \; .$$

Section Oliver Knill, knill@math.harvard.edu, Harvard College, Fall 2022

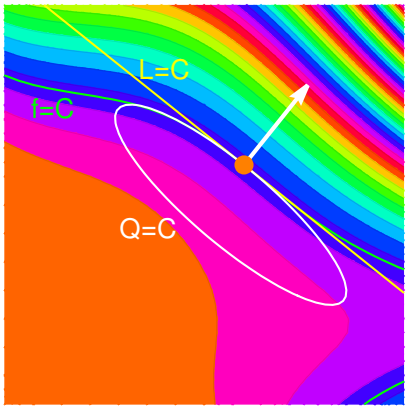

FIGURE 1. The linear and quadratic approximations $L(x, y)$ and $Q(x, y)$ in the contour picture. We see the contours of f and the contour of L and Q passing through the points. We see also that the gradient is perpendicular to all these three curves at (x_0, y_0).

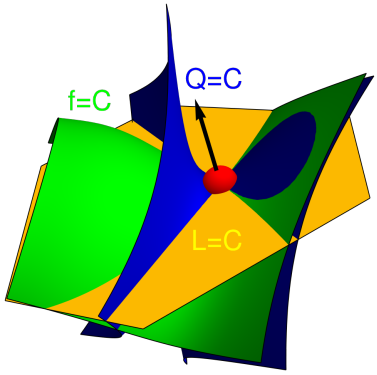

FIGURE 2. The linear and quadratic approximations $L(x, y)$ and $Q(x, y)$ are seen as graphs. They are both tangent to the graph.

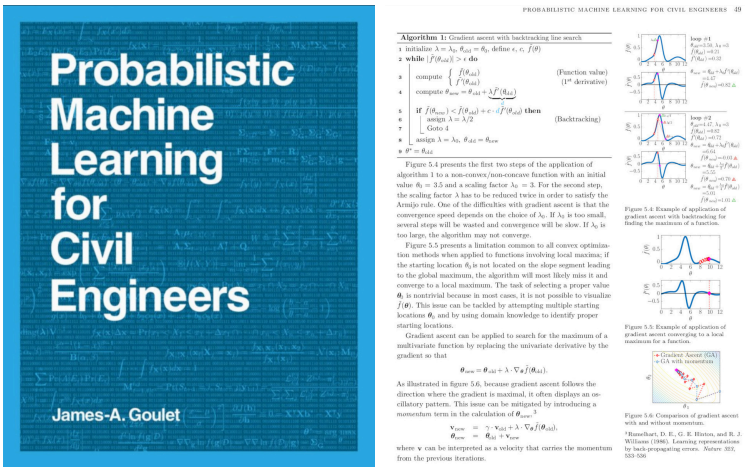

FIGURE 3. The gradient is important in machine learning. It can be used there to learn. Learn it to use it!