3/5/2003, LINEAR APPROXIMATION — Math 21a, O. Knill

Homework for Friday:
Section 11.4: 4, 12, 28, 30, 40

RECALL PARTIAL DERIVATIVES. $f_x(x,y) = \partial f/\partial x$ is defined as the derivative with respect to x, when y is treated as a constant.

Because $\frac{f(x+dx,y)-f(x,y)}{dx} \to f_x(x,y)$ for $dx \to 0$, we have $f_x(x,y)dx \sim f(x+dx,y) - f(x,y)$ and $f_x(x,y)$ is the **rate of change** of f at the point (x,y) **in the x direction**.

EXAMPLE. If $f(x,y) = \sin(xy)$, then $f_x(x,y) = y\cos(xy)$ is again a function of two variables. At a specific point, like $(\pi, 1)$, we have $f_x(\pi,1) = -1$.

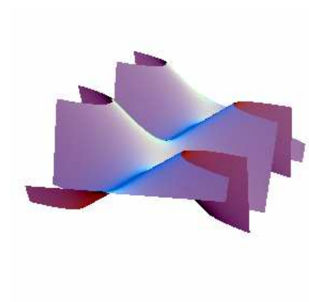
The estimate $f(x+dx,y) \sim f(x,y) + f_x(x,y)dx$ is a **linear approximation** of f. In the above example, if we take $dx = 0.01$, then $f(\pi + 0.01, 1) = f(\pi,1) + 0.01 f'(\pi,1) = -0.00999983$ is only by $2 \cdot 10^{-7}$ away from the correct answer. Today, we will see how to do such approximations in higher dimensions.

TAKING PARTIAL DERIVATIVES OF IMPLICIT EQUATIONS. Let $f(x,y,z) = x^5 - 3y^5 + z^5 + 3xyz^3 + z = 0$. At most points, we can in principle solve for z so that $z = z(x,y)$ nearby. What is $\partial z/\partial x$?

Surprisingly, the answer can be given without knowing $z(x,y)$: $f_x(x,y,z(x,y)) = 5x^4 - 3y^5 + 5z^4 z_x + 3yz^3 + 9xyz^2 z_x + z_x = 0$. Solving for z_x gives $z_x = \frac{5x^4 - 3y^5 + yz^3}{5z^4 + 9xyz^2 - 1}$.

An example, where we can find the implicit function is $f(x,y) = x^2 + y^2 - 1 = 0$ which defines a circle. What is $y_x = \frac{\partial y}{\partial x}$?
$2x + 2yy_x = 0$ gives $y_x = -2x/2y = -x/y$. We can also compute $y(x) = \sqrt{1-x^2}$ and obtain $y_x(x) = (1/2)(-2x)/\sqrt{1-x^2} = -x/\sqrt{1-x^2} = -x/y$, the same result.

LINEAR APPROXIMATION. **1D:** The linear approximation of a function $f(x)$ at a point x_0 is the linear function $g(x) = f(x_0) + f'(x_0)(x - x_0)$.

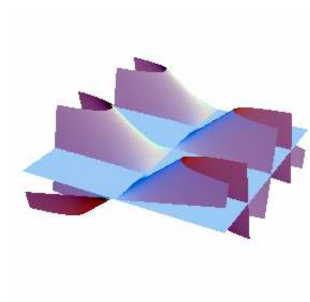


2D: A function $f(x,y)$ is approximated at (x_0, y_0) by $g(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(y-y_0)$.

The level curves of g is tangent to the level curve of f at (x_0, y_0).

3D: A function $f(x,y,z)$ is approximated at (x_0,y_0,z_0) by
$g(x,y,z) = f(x_0,y_0,z_0) + f_x(x_0,y_0,z_0)(x-x_0) + f_y(x_0,y_0,z_0)(y-y_0) + f_z(x_0,y_0,z_0)(z-z_0)$.
The level surfaces of g are planes which are tangent to the level surface of f at the point (x_0,y_0,z_0).

To write this in a more compact way, introduce the vector $\nabla f = (f_x, f_y)$ in 2D or $\nabla f = (f_x, f_y, f_z)$ in 3D called the **gradient**. It will play an important role later. In 1D, $\nabla f(x) = f'(x)$ is the usual derivative.



The function

$g(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$

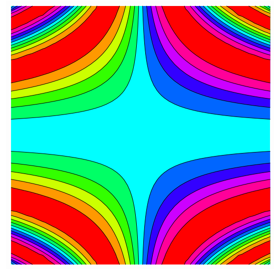
in each of the above cases is called the **linear approximation** of f at the point $\vec{x}_0$.

JUSTIFYING THE LINEAR APPROXIMATION.

1) If we fix $y = y_0$ and vary x, then $f(x, y_0) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0)$ is the correct linear approximation of the function (fixing y leads to a function of one variable). Similarly, if we fix $x = x_0$ and vary y, then $f(x_0, y) = f(x_0, y_0) + f_y(x_0, y_0)(y - y_0)$. So along two directions, the linear approximations are the best. Two directions determine the plane already.

2) We will see next week an other proof, which uses that $\nabla f(\vec{x}_0)$ is orthogonal to the level curve at $\vec{x}_0$. Because $\vec{n} = \nabla f(\vec{x}_0)$ is orthogonal to the plane $\vec{n} \cdot (\vec{x} - \vec{x_0}) = d$, the fact will follow again.

EXAMPLE (2D) Find the linear approximation of the function $f(x, y) = \sin(\pi x y^2)$ at the point $(1, 1)$. The gradient is $\nabla f(x, y) = (\pi y^2 \cos(\pi x y^2), 2y\pi \cos(\pi x y^2))$. At the point $(1, 1)$, we have $\nabla f(1, 1) = (\pi \cos(\pi), 2\pi \cos(\pi)) = (-\pi, 2\pi)$. The linear function approximating f is $g(x, y) = f(1, 1) + \nabla F(1, 1) \cdot (x - 1, y - 1)$ $= 0 - \pi(x - 1) - 2\pi(y - 1) = -\pi x - 2\pi y + 3\pi$. The level curves of G are the lines $x + 2y = const$. The line which passes through $(1, 1)$ satisfies $x + 2y = 3$.



Application: $-0.00943407 = f(1+0.01, 1+0.01) \sim g(1+0.01, 1+0.01) = -\pi 0.01 - 2\pi 0.01 + 3\pi = -0.00942478$.

EXAMPLE (2D) Find the linear approximation to $f(x, y, z) = xy + yz + zx$ at the point $(1, 1, 1)$.

We have $f(1, 1, 1) = 3$, $\nabla f(x, y, z) = (y+z, x+z, y+x)$, $\nabla f(1, 1, 1) = (2, 2, 2)$. Therefore $g(x, y, z) = f(1, 1, 1) + (2, 2, 2) \cdot (x - 1, y - 1, z - 1) = 3 + 2(x - 1) + 2(y - 1) + 2(z - 1) = 2x + 2y + 2z - 3$.

EXAMPLE (3D). Use the best linear approximation to $f(x, y, z) = e^x \sqrt{y} z$ to estimate the value of f at the point $(0.01, 24.8, 1.02)$.

Solution. Take $(x_0, y_0, z_0) = (0, 25, 1)$, where $f(x_0, y_0, z_0) = 5$. The gradient is $\nabla f(x, y, z) = (e^x \sqrt{y} z, e^x z/(2\sqrt{y}), e^x \sqrt{y})$. At the point $(x_0, y_0, z_0) = (0, 25, 1)$ the gradient is the vector $(5, 1/10, 5)$. The linear approximation is $g(x, y, z) = f(x_0, y_0, z_0) + \nabla f(x_0, y_0, z_0)(x - x_0, y - y_0, z - z_0) = 5 + (5, 1/10, 5)(x - 0, y - 25, z - 1) = 5x + y/10 + 5z - 2.5$. We can approximate $f(0.01, 24.8, 1.02)$ by $5 + (5, 1/10, 5) \cdot (0.01, -0.2, 0.02) = 5 + 0.05 - 0.02 + 0.10 = 5.13$. The actual value is $f(0.01, 24.8, 1.02) = 5.1306$, very close to the estimate.

TANGENT PLANES/CURVES. Geometrically, linearisation means finding the **tangent plane** to a level surface $f(x, y, z) = c$ or to find the **tangent curve** to $f(x, y) = c$. If the linear function $g(x, y, z)$ approximates $f(x, y, z)$ at a point $P = (x, y, z)$, then the level curve of f at P and the level curve of g at P are the tangent. Indeed, $\nabla g(x, y, z)$ is equal to $\nabla f(x, y, z)$ at (x, y, z).

The parameters of the **tangent line** $ax + by = d$ to the curve $f(x, y) = C$ at the point (x_0, y_0) are $(a, b) = \nabla f(x_0, y_0)$ and $d = ax_0 + by_0$.

The parameters of the **tangent plane** $ax + by + cz = d$ to the surface $f(x, y, z) = C$ at the point (x_0, y_0, z_0) are $(a, b, c) = \nabla f(x_0, y_0, z_0)$ and $d = ax_0 + by_0 + cz_0$.

EXAMPLE (2D). Find the tangent line to the hyperbola $x^2 - 2y^2 = 1$ at the point $(3, 2)$.
If $f(x, y) = x^2 - y^2$, then $\nabla f(x, y) = (2x, -2y)$ and $\nabla f(3, 2) = (6, -4)$. The linear approximation is $g(x, y) = 1 + (6, -4) \cdot (x - 3, y - 2)$. The line is $6x - 4y = 10$.

EXAMPLE (3D). Find the tangent plane to the two-sheeted hyperboloid $x^2 + y^2 - z^2 = -1$ at the point $(x_0, y_0, z_0) = (2, 2, 3)$. The hyperboloid is a level-surface to the function $F(x, y, z) = x^2 + y^2 - z^2$, whose gradient is $(2x, 2y, -2z)$. At the point $(2, 2, 3)$, the gradient is $(4, 4, 6)$. The equation of the plane is therefore $4x + 4y - 6z = -2$.

DIFFERENTIABILITY. If $f(\vec{x})$ is a function of several variables, we say it is **differentiable** at $\vec{x}_0$ if all partial derivatives exist there and the linear approximation $g(\vec{x})$ satisfies $(f(\vec{x}) - g(\vec{x}))/|\vec{x} - \vec{x}_0| \to 0$ for $\vec{x} \to \vec{x}_0$.

INTERPRETATION OF DIFFERENTIABILITY. The condition means that the distance of the tangent plane at $\vec{x}_0$ to the surface $f(\vec{x}) = 0$ goes to zero faster than $|x - x_0|$.

PARAMETRIC SURFACES. A parametric surface was given by a map $X(u,v) = (f(u,v), g(u,v), h(u,v))$. If we fix v_0, then $u \mapsto X(u, v_0)$ is a grid curve on the plane and $X_u(u_0, v_0)$ is the velocity of this curve at the point $X(u_0, v_0)$. It is tangent to the curve and therefore tangent to the surface. Similarly, $X_v(u_0, v_0)$ is tangent to the surface. The vector $(a,b,c) = n = X_u \times X_v(u_0, v_0)$ is orthogonal to the surface at the point $X(u_0, v_0)$. The plane $ax+by+cz = d$ is the tangent plane at the point $(x_0, y_0, z_0) = X(u_0, v_0)$. The constant d is $ax_0+by_0+cz_0$ as usual.

EXAMPLE. $X(u,v) = (u, v, u^2 + v^2)$ describes a paraboloid $f(x,y,z) = x^2 + y^2 - z = 0$. Let us compute the tangent plane at $(1,2,5) = X(1,2)$ in two ways.

1) As a parametrized surface, we have $X_u = (1, 0, 2u), X_v = (0, 1, 2v)$ and $X_u \times X_v(u,v) = (-2u, -2v, 1)$ so that $X_u \times X_v(1,2) = (-2, -4, 1)$. The plane is $-2x - 4y + z = d = -21 - 42 + 5 = -5$.

2) We can also compute $\nabla f(x,y,z) = (2x, 2y, -1)$ and $\nabla f(1,2,5) = (2,4,-1)$. The plane is $2x + 4y - z = d = 21 + 42 - 5 = 5$, the same result as in 1).

PHYSICAL LAWS. Many physical laws are in fact linear approximations to more complicated laws. One could say that a large fraction of physics consists of understand nature with linear laws.

Linear stability analysis. In physics, complicated situations can occur. Usually, many unknown parameters are present and the only way to analyze the situation theoretically is to assume that things depend linearly on these parameters. The analysis of the linear situation allows then to predict for example the stability of the system with respect to perturbations. Sometimes, the stability of the linearized system will imply the stability of the perturbation.

Error analysis. Error analysis is based on linear approximation. Assume, you make a measurement of a function $F(a,b,c)$, where a,b,c are parameters. Assume, you know the numbers a,b,c up to accuracy ϵ. How precise do you know the values $F(a,b,c)$? Because $F(a_0 + \epsilon_a, b_0 + \epsilon_b, c_0 + \epsilon_c)$ is about $F(a_0, b_0, c_0) + \nabla F(a_0, b_0, c_0) \cdot (\epsilon_a, \epsilon_b, \epsilon_c)$, the answer is that we know F up to accuracy $|\nabla F(a_0, b_0, c_0)|\epsilon$.

Power laws. Some laws in physics are given by functions of the form $f(x,y) = x^\alpha y^\beta$. An example is the Cobb-Douglas formula in economics. Such dependence on x or y is called **power law behavior**. If we consider $F = \log(f)$, and introduce $a = \log(x), b = \log(y)$, then this becomes $F(a,b) = \log(f(x,y)) = a\alpha + b\beta$. Power laws become linear laws in a logarithmic scale. But they usually are linear approximations to more complicated nonlinear relations.

EXAMPLES OF LINEARIZED LAWS.

Electronics. If we apply a voltage difference U at the ends of a resistor R, then a current I flows. The relation $U = RI$ is called **Ohms law**. In logarithmic coordinates $\log(U) = \log(R) + \log(I)$, this is a linear law. In reality, the relation between current, voltage and resistance is more complicated. For example, if the resistor heats up, then its characteristics begin to change. Nonlinear resistors are used for example in synthesizers or in radars. While Ohm's law works **extremely well**, the nonlinear behavior can have important consequences for example to stabilize systems or to protect equipment against over-voltages.

Thermodynamics. If $l(T)$ is the length of an object with temperature T, then $l(T) = l(T_0) + c(T - T_0)$, where the expansion coefficient c depends on the material. (Trick question: What happens if you heat a ring, does the inner ring become smaller or bigger?). The volume of a hot air balloon and therefore its lift capacity grows like $c(T - T_0)^3$. The law of expansion is only an approximation.

Oceanography For oceanographers, it is important to know the water density $\rho(T,S)$ in dependence on the **temperature** T (Kelvin) and the **salinity** S (psu). If we would include the pressure P (Bar), then we had a function $\rho(T,S,P)$ of three variables. Near a specific point (T,S,P) the density can be approximated by a linear function giving a law which is precise enough.

Engineering Hooks law tells that the force of a spring is proportional to the length with which it is pulled: $F(l) = c(l - l_0)$, where l_0 is the length when the spring is relaxed. This allows to measure weights or to cushion shocks. However, this law is only good in a certain range. If the spring is pulled too strongly, then more force is needed. Such a nonlinear behavior is needed for example in shock absorbers.

Mechanics. For small amplitudes, the pendulum motion $\ddot{x} = -g\sin(x)$ can be approximated by $\ddot{x} = -gx$, the harmonic oscillator. Nonlinear (partial) differential equations like $u_{xx} + u_{yy} + u_{zz} = F(x,y,z)$ are often approximated by linear differential equations.

Cartography. It was well known already to the Greeks that we live on a sphere. On a sphere a triangle however the sum of its angles adds up to more than 180 degrees and every straight line (great circles) crosses every other line at least twice. Despite this, a city map can perfectly assume that the coordinate system is Cartesian. When drawing a plan of a house, an architect can assume that the house stands on a plane (the level curve of the linearisation $G(x,y)$ of $F(x,y)$ defining the surface of the earth.)

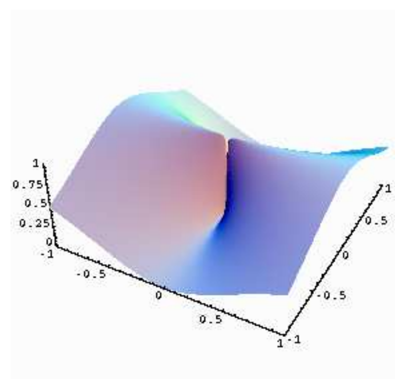
Relativity. Newton's law tells that $r''(t)$, the acceleration of a particle is proportional to the force F which acts on the mass point: $r''(t) = F/m$. For a constant force and zero initial velocity this implies $r'(t) = tF/m$. This law can not apply for all times, because we can not reach the speed of light with a massive body. In special relativity, the Newton axiom is replaced with $d/dt(r'(t)m(t)) = F$, where the mass $m(t)$ depends on the velocity. This gives $v(t) = (tF/m_0)\frac{1}{\sqrt{1+F^2t^2/(c^2m_0^2)}}$. Linearisation at $t = 0$ produces the classical law $v(t) = tF/m_0$.

Economics. The mathematician Charles W. Cobb and the economist Paul H. Douglas found in 1928 empirically a formula $F(L,K) = bL^\alpha K^\beta$ giving the total production F of an economic system as a function of the amount of labor L and the capital investment K. This is a linear law in logarithmic coordinates. The formula actually had been found by linear fit of empirical data. In general, the production depends in a more complicated way on labor and capital investment. For example, with increase of labor and investment, logistic constraints will become relevant.

Chemistry. The ideal gas law $PV = RT$ relates the pressure, the volume and the temperature of an ideal gas using a constant R called the Avogadro number. This law $T = f(P,V)$ is linear in logarithmic scales. This law is only an approximation and has to be replaced by the van der Waals law, which takes into account the molecular interactions as well as the volume of the molecules.

SMOOTHNESS. Linear approximation requires that the function is smooth enough. Points, where linear approximations are not possible can appear. Examples in physics are **phase transitions**. Near such points, the linear approximation usually is invalid: the function is not differentiable there.

An example of a function for which linear approximation does not work is $\boxed{f(x,y) = x^2/(x^2+y^2)}$. Because $f(x,y) = 0$ for $x = 0$ and $f(x,y) = 1$ for $y = 0$, we have $f_x(x,y) = 0$ and $f_y(x,y) = 0$. Therefore, the gradient is $(0,0)$. However $f(x,x) = x^2/(2x^2) = 1/2$ for every x. The approximation $f(x,y) = f(0,0) + \nabla f(0,0) \cdot (x,y)$ is not valid.



GRADIENT. If $f(x,y)$ is a function of two variables it has the gradient $\nabla f(x,y) = (f_x(x,y), f_y(x,y))$.

If $f(x,y,z)$ is a function of three variables, then $\nabla f(x,y,z) = (f_x(x,y,z), f_y(x,y,z), f_z(x,y,z))$ is called the **gradient** of f.

The symbol ∇ is called **Nabla**.

NORMAL. Important fact: The gradient $\nabla f(x,y)$ is normal to the level curve $f(x,y) = c$ and the gradient $\nabla f(x,y,z)$ is normal to the level surface $f(x,y,z)$. (For example, the gradient of $f(x,y,z) = x^2 + y^2 - z^2$ at a point (x,y,z) is $(2x, 2y, -2z)$.) We will see a proof of the fact next week.