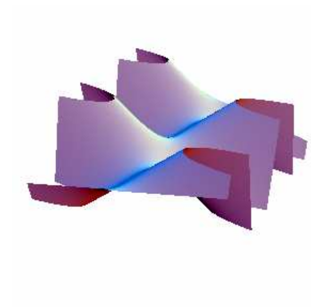
Lecture 12: 3/8/2004, LINEAR APPROXIMATION **Math 21a, O. Knill**

HOMEWORK: 11.6: 4,12,18,40 (for Friday)

LINEAR APPROXIMATION.

1D: The **linear approximation** of a function $f(x)$ at a point x_0 is the linear function $L(x) = f(x_0) + f'(x_0)(x - x_0)$. The graph of L is tangent to the graph of f.



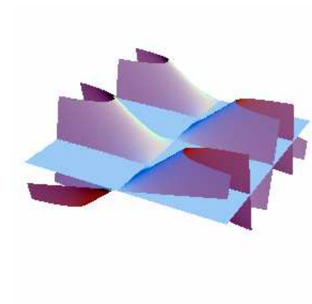
2D: The **linear approximation** of a function $f(x, y)$ at (x_0, y_0) is

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(y - y_0)$$

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

3D: The **linear approximation** of a function $f(x, y, z)$ at (x_0, y_0, z_0) by

$$L(x, y, z) = f(x_0, y_0, z_0) + f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0).$$



The level surface of L is tangent to the level surface of f at (x_0, y_0, z_0).

In vector form, the linearization can be written as

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

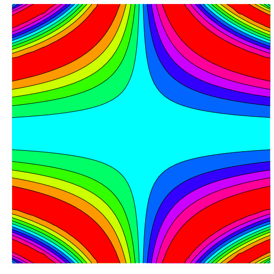
JUSTIFYING THE LINEAR APPROXIMATION. 3 ways to see it:

1) We know the tangent plane to $g(x, y, z) = z - f(x, y)$ at $(x_0, y_0, z + 0 = f(x_0, z_0)$ is $-f_x x - f_y y + z = -f_x x_0 - f_y y_0 + z_0$. This can be read as $z = z_0 + +f_x(x - x_0) + f_y(y - y_0)$. Calling the right hand side $L(x, y)$ shows that the graph of L is tangent to the graph of f at (x_0, y_0).

2) The higher dimensional case can be reduced to the one dimensional case: if $y = y_0$ is fixed and x, then $f(x, y_0) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0)$ is the linear approximation of the function. Similarly, if $x = x_0$ is fixed y is the single variable, 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. Together we get the approximation for $f(x, y)$.

3) An other justification is 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 $\vec{n} \cdot (\vec{x} - \vec{x}_0) = d$ also, the graphs of $L(x, y)$ and $f(x, y)$ have the same normal vector at $(x_0, y_0, f(x_0, y_0))$.

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 $L(x, y) = f(1, 1) + \nabla f(1, 1) \cdot (x - 1, y - 1)$ $= 0 - \pi(x - 1) - 2\pi(y - 1) = -\pi x - 2\pi y + 3\pi$. 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 (3D) 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 $L(x, y, z) = f(1, 1, 1) + (2, 2, 2) \cdot (x - 1, y - 1, z - 1) = 3 + 2(x - 1) + 2(y - 1) + 2(z - 1) = 2x + 2y + 2z - 3$.

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 $L(x,y,z) = f(x_0,y_0,z_0) + \nabla f(x_0,y_0,z_0)(x - x_0, y - y_0, z - z_0) = 5 + (5, 1/10, 5)(x - 0, y - 25, z - 1) = 5x + y/10 + 5z - 2.5$. 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.

SECOND DERIVATIVE.
If $f(x,y)$ is a function of two variables, then the matrix $f''(x,y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$ is called the **second derivative** or the **Hessian** of f.

For functions of three variables, the Hessian is the 3x3 matrix $f''(x,y,z) = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}$ Because for smooth functions, $f_{xy} = f_{yx}, f_{yz} = f_{zy}$, the matrix f'' is **symmetric** (a reflection at the diagonal leaves it invariant).

QUADRATIC APPROXIMATION. (informal)
If F is a function of several variables $\vec{x}$ and $\vec{x}_0$ is a point, then

$$Q(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0)(\vec{x} - \vec{x}_0) + [f''(\vec{x}_0)(\vec{x} - \vec{x}_0)] \cdot (\vec{x} - \vec{x}_0)/2$$

is called the **quadratic approximation** of $\vec{x}$.

It generalizes the quadratic approximation $L(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)(x - x_0)^2/2$ of a function of one variables.

EXAMPLE. If the height of a hilly region is given by $f(x,y) = 4000 - \sin(x^2 + y^2)$, find the quadratic approximation of F at $(0,0)$.
$\nabla f(x,y) = (2x, 2y)\cos(x^2 + y^2)$ so that $\nabla f(0,0) = (0,0)$. The linear approximation of F at $(0,0)$ is $G(x,y) = f(0,0) = 4000$. The graph of G is a plane tangent to the graph of F.

$$f''(x,y) = \cos(x^2 + y^2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

so that

$$f''(0,0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} .$$

The quadratic approximation at $(0,0)$ is $Q(x,y) = 4000 - x^2 - y^2$. The graph of F is an inverted paraboloid.

EXAMPLE: REDUCED VAN DER WAALS LAW $T(p,V) = (p + 3/V^2)(3V - 1)/8$ Find the quadratic approximation of T at $(p,V) = (1,1)$. We had $T_p(p,V) = (3V - 1)/8$ and $T_V(p,V) = 3p/8 - (9/16)1/V^2 - 3/(4V^3)$ and $\nabla T(1,1) = (1/4, 0)$. Now, $T_{pp}(p,V) = 0$, $T_{pV} = 3/8$, $T_{Vp} = 3/8$ and $T_{VV} = (9/8)1/V^3 + (9/4)1/V^4$ so that $T''(1,1) = \begin{bmatrix} 0 & 3/8 \\ 3/8 & 27/8 \end{bmatrix}$. The quadratic approximation at $T = (1,1)$ is $Q(p,T) = T(1,1) + T'(1,1) \cdot (p - 1, T - 1) + [T''(1,1)(p - 1, T - 1)] \cdot (p - 1, T - 1)/2$.

ERROR OF APPROXIMATION. It follows from Taylors theorem that the error $f(x,y) - L(x,y)$ in a region R near (x_0, y_0) is smaller or equal to $M(|x - x_0| + |y - y_0|)^2/2$, where M is the maximal value of all the matrix entries $f''(x,y)$ in that region R.

TOTAL DIFFERENTIAL. Aiming to estimate the change $\Delta f = f(x,y) - f(x_0,y_0)$ of f for points $(x,y) = (x_0,y_0) + (\Delta x, \Delta y)$ near (x_0, y_0), we can estimate it with the linear approximation which is $L(\Delta x, \Delta y) = f_x(x_0,y_0)\Delta x + f_y \Delta y$. In an old-fashioned notation, one writes also $df = f_x dx + f_y dy$ and calls df the **total differential**. One can **totally avoid** the notation of the **total differential**.

RELATIVE AND ABSOLUTE CHANGE. The **relative change** Δf depends on the magnitude of f. One also defines $\Delta f(x_0,y_0)/f(x_0,y_0)$, the **relative change** of f. A good estimate is $\Delta L(x_0,y_0)/L(x_0,y_0)$.

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

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.

ENGENEERING.

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 linearization $G(x, y)$ 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 $v(t) = (tF/m_0)\frac{1}{\sqrt{1+F^2t^2/(c^2m_0^2)}}$. Linearization 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.