



Kantorovitch's theorem

Canonical name	KantorovitchsTheorem
Date of creation	2013-03-22 11:58:09
Last modified on	2013-03-22 11:58:09
Owner	stevecheng (10074)
Last modified by	stevecheng (10074)
Numerical id	25
Author	stevecheng (10074)
Entry type	Theorem
Classification	msc 49K10
Synonym	Kantorovitch inequality
Related topic	LipschitzCondition
Related topic	NewtonsMethod
Related topic	Superconvergence

Let \mathbf{a}_0 be a point in \mathbb{R}^n , U an open neighborhood of \mathbf{a}_0 in \mathbb{R}^n and $\mathbf{f}: U \rightarrow \mathbb{R}^n$ a differentiable mapping, with its derivative $[\mathbf{Df}(\mathbf{a}_0)]$ invertible. Define

$$\mathbf{h}_0 = -[\mathbf{Df}(\mathbf{a}_0)]^{-1}\mathbf{f}(\mathbf{a}_0), \mathbf{a}_1 = \mathbf{a}_0 + \mathbf{h}_0, U_0 = \{\mathbf{x} \mid |\mathbf{x} - \mathbf{a}_1| \leq |\mathbf{h}_0|\}.$$

If $U_0 \subset U$ and the derivative $[\mathbf{Df}(\mathbf{x})]$ satisfies the <http://planetmath.org/node/765>Lipschitz condition

$$|[\mathbf{Df}(\mathbf{u}_1)] - [\mathbf{Df}(\mathbf{u}_2)]| \leq M|\mathbf{u}_1 - \mathbf{u}_2|$$

for all points $\mathbf{u}_1, \mathbf{u}_2 \in U_0$, and if the inequality

$$|\mathbf{f}(\mathbf{a}_0)| |[\mathbf{Df}(\mathbf{a}_0)]^{-1}|^2 M \leq \frac{1}{2}$$

is satisfied, the equation $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ has a unique solution in U_0 , and Newton's method with initial guess \mathbf{a}_0 converges to it. If we replace \leq with $<$, then it can be shown that Newton's method <http://planetmath.org/node/793>superconverges! If you want an even stronger version, one can replace $|\dots|$ with the norm $\|\dots\|$.

Logic behind the theorem:

Let's look at the useful part of the theorem:

$$|\mathbf{f}(\mathbf{a}_0)| |[\mathbf{Df}(\mathbf{a}_0)]^{-1}|^2 M \leq \frac{1}{2}.$$

It is a product of three distinct properties of your function such that the product is less than or equal to a certain number, or bound. If we call the product R , then it says that \mathbf{a}_0 must be within a ball of radius R . It also says that the solution \mathbf{x} is within this same ball. How was this ball defined?

The first term, $|\mathbf{f}(\mathbf{a}_0)|$, is a measure of how far the function is from the domain; in the Cartesian plane, it would be how far the function is from the x-axis. Of course, if we're solving for $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, we want this value to be small, because it means we're closer to the axis. However a function can be annoyingly close to the axis, and yet just happily curve away from the axis. Thus we need more.

The second term, $|[\mathbf{Df}(\mathbf{a}_0)]^{-1}|^2$ is a little more difficult. This is obviously a measure of how fast the function is changing with respect to the domain (x-axis in the plane). The larger the derivative, the faster it's approaching wherever it's going (hopefully the axis). Thus, we take the inverse of it, since we want this product to be *less than* a number. Why it's squared though, is

because it is the denominator where a product of two terms of like units is the numerator. Thus to conserve units with the numerator, it is multiplied by itself. Combined with the first term, this also seems to be enough, but what if the derivative changes sharply, but it changes the wrong way?

The third term is the Lipschitz ratio M . This measures sharp changes in the first derivative, so we can be sure that if this is small, that the function won't try to curve away from our goal on us too sharply.

By the way, the number $\frac{1}{2}$ is unitless, so all the units on the left side cancel. Checking units is essential in applications, such as physics and engineering, where Newton's method is used.