Mathematical Modelling for Computer Networks- Part I **Spring 2013 (Period 4)**

Math Appendix 2: Calculus

Lecturers: Laila Daniel and Krishnan Narayanan *Date:11th March 2013*

To see a World in a Grain of Sand
And a Heaven in a Wild Flower
Hold Infinity in the palm of your hand
And Eternity in an hour
- William Blake

Brief review of some notions from Calculus

Some preliminaries about the real line $\mathbb{R}$
Recall that the set of real numbers $\mathbb{R}$ is the playground of basic calculus. There is a key property of the real line which confers its preeminent status in calculus - the property known as *completeness* of the set of reals. This is also sometimes called the *continuity property* of the set of real numbers. This property can be stated in various equivalent ways. Two such equivalent characterizations are the following.

Completeness property of $\mathbb{R}$
C1: Every set of real numbers that is bounded above has a least upper bound (lub).
C2: Every bounded monotone sequence converges.

Informally, the completeness property of $\mathbb{R}$ says that the number line whose points can be thought of as real numbers is without 'holes'. As we shall see shortly, the set of rational numbers $\mathbb{Q}$ lacks this crucial completeness property of $\mathbb{R}$ which makes $\mathbb{Q}$ completely unsuitable to build a theory of calculus.

The significance of the completeness of the set of real numbers, $\mathbb{R}$, is exemplified by the next two propositions.
For any positive real number x, the n-th root $\sqrt[n]{x}$ exists for every natural number n. For any two positive real numbers x and y, the square root of the product xy is the product of the square roots of x and y: $\sqrt{xy} = \sqrt{x}\sqrt{y}$.
Dedekind, who invented the theory of completeness of the real numbers in the late 19th century, rigorously established the above result and remarked that as a consequence we know that $\sqrt{2.3} = \sqrt{6} = \sqrt{2}\sqrt{3}$ after almost two thousand years since the Greek logician Eudoxis began the study of such questions.

Example:
Consider the set X of rational numbers whose square is less than 2. We write this set X as

$$X \doteq \{x \in \mathbb{Q} \,|\, x^2 < 2\}$$

Clearly the set is bounded above. For example, 2 is an upper bound for X but the set X does not have least upper bound (lub) in $\mathbb{Q}$, as $\sqrt{2}$ is a well-known to be an irrational number. However, the set X has $\sqrt{2}$ as lub in $\mathbb{R}$. We observe that the set X has rational numbers which lie arbitrarily close to $\sqrt{2}$ (This property is called '$\mathbb{Q}$ is *dense* in $\mathbb{R}$'). So the set of rationals has 'holes' corresponding to the irrational numbers and the set of real numbers can be regarded as obtained from $\mathbb{Q}$ by filling these ' holes'. This is the geometric significance of the completeness property of the set of real numbers.

Open interval and closed interval

The *open interval* $(a,\ b)$ denotes the set of real numbers that lie between a and b with both the end points a and b excluded.
The *closed interval* $[a,\ b]$ denotes the set of real numbers that lie between a and b with both the end points a and b included in the set.

Symbolically we write this as follows.

$$(a\,b) = \{x \in \mathbb{R} \mid a\ <\ x\ < b\}$$
$$[a\,b] = \{x \in \mathbb{R} \mid a\ \leq\ x\ \geq b\}$$

***lub* and *glb* as operators on sets**
Let X be a subset of $\mathbb{R}$, then $lub(X)$ gives the least upper bound of the set X in $\mathbb{R}$. $lub(X)$ is also denoted by $sup(X)$. Here *sup* stands for *supremum*.
Similarly $glb(X)$ gives the greatest lower bound of the set X in $\mathbb{R}$. $glb(X)$ is also denoted by $inf(X)$. Here *inf* stands for *infimum*.
Notice that the operators *lub* and *glb* are *duals* of each other. This means

$$lub(X) = -glb(-X) \quad \text{whenever they exist}$$

By convention when the set X is not bounded from above, $lub(X)$ is denoted by $+\infty$. Similarly when X is not bounded below, $glb(X)$ is denoted by $-\infty$.

Example:
$lub(0,\ 1) = 1$ which does not lie in the set $(0,\ 1)$.
$glb(0,\ 1) = 0$ which does not lie in the set $(0,\ 1)$.
So in general $lub(X)$ and $glb(X)$ do not necessarily lie in the set (X), whenever these operators are defined on the set.
Of course
$lub[0,\ 1] = 1$ which lies in the set $[0,\ 1]$.
Similarly
$glb[0,\ 1] = 0$ which lies in the set $[0,\ 1]$.

Local/global maxima and minima and stationary points
Consider a function defined over an open interval (a, b) and let $\xi \in (a, b)$. Then f is a *local maximum* if

$$f(x) \leq f(\xi) \text{ for all values of } x \text{ in some open interval } I \text{ containing } \xi$$

Analogously we define *local minima* at the point ξ by a similar definition as above except that the inequality in the definition above is reversed.

Roughly speaking, at a local maximum, the graph of the function resembles a 'little hill' whereas at a local minimum it resembles a 'small valley'.

If a function has the maximum value of $f(\xi)$ at ξ in the whole interval (a, b) then f has a *global maximum* at ξ. If the function has a global maximum at ξ then obviously it has a

local maximum at ξ as well, but the converse is not true.
Analogously we can define the *global minimum* of a function.

The next result describes a necessary condition for a differentiable function f in the interval (a,b) to have either a local maximum or local minimum at the point ξ . As it is fairly apparent that a tangent to the curve at the point ξ of a local maxima or minima is horizontal, necessarily the derivative of the function vanishes (derivative equals zero) at ξ. Note that the converse is not true; i.e., at a point where a derivative vanishes need not be either a local maxima or a local minima.

If the function is twice differentiable and the second derivative of the function at the point ξ is greater than zero, i.e. $D^2 f(\xi) > 0$, then the function has a local minima at ξ. Similarly when the second derivative of the function at the point ξ is less than zero, i.e. $D^2 f(\xi) < 0$, then the function has a local maxima at ξ.

Stationary point
A point where the derivative vanishes is a *stationary point*. Not all stationary points give rise to a local maximum or local minimum. In this case the stationary point may be a point of *inflection* where the function changes from a concave function to a convex function or vice versa.
A stationary point need not be any of the above three points (local maxima, local minima or inflection point). This is a pathological case where the function may oscillate with decaying amplitude going to zero at the point where the derivative vanishes.

Taylor series
Consider a function f differentiable as often as we want in an open interval containing a point ξ. Then the Taylor expansion of f at this point is given by a polynomial expansion around the point ξ whose coefficients involve various derivatives of the function at the point ξ. We denote the series expansion by $Tf(x)$

$$Tf(x) = \sum_{n=0}^{\infty} \frac{(x-\xi)^n}{n!} f^{(n)}(\xi) = f(\xi) + \frac{(x-\xi)}{1!} f'(\xi) + \frac{(x-\xi)^2}{2!} f''(\xi) + \ldots$$

Taylor series is used to approximate a function in viscinity of a chosen point.

Gradient The notion of *gradient* of a function of one or more variables can be illustrated by a simple example.
Let $f(x, y) \doteq x^2 + y^2$. The contours of level curves are given by $x^2 + y^2 = k$ for various values of $k \in \mathbb{R}^+$. The level curves correspond to the family of concentric circles around the origin.
Along which *direction* one should move at any point so that the value of the function f increases most rapidly? And how steep is the increase? The answer is clear in this case, move along the radial direction which can be formally stated as moving along the direction of the outward normal which is perpendicular to the tangent drawn at each point (x, y) to the curve $f(x, y) = x^2 + y^2$. However, the quantity to assign to the steepness is not so intuitively evident. The notion of the gradient captures this intuition by giving a complete answer to the question.

The function $f(x, y)$ can be visualised as the altitude corresponding to the point (x, y) in the plane. The contours identify the points (x, y) in the plane which correspond to the same height (similar idea appears as isobars, isotherms etc which correspond to contours of equal pressure, temperature etc).

The *gradient* of the function f denoted by ∇f gives both the magnitude and direction of the maximum rate of change of the function at the point (x, y). The gradient at the point (x, y) can be computed by evaluating the partial derivatives of the function at that point.

$$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$

So for $f(x, y) = x^2 + y^2$, $\nabla f(x, y) = (2x, 2y)$ which corresponds to the vector $2x\mathbf{i} + 2y\mathbf{j}$, where $\mathbf{i}$ and $\mathbf{j}$ are unit vectors in x and y direction respectively.

Let $D_u f(\xi)$ denote the *directional derivative* of the function f in a direction given by the unit vector u at a point ξ .
$D_u f(\xi)$ is obtained by taking the component of the gradient at the point ξ along the direction $\mathbf{u}$. This can be expressed using the inner product as follows.

$$D_u f(\xi) \;= (\nabla f(\xi), u)$$

Using the Cauchy-Schwartz inequality for the inner product which says $|\,(x\,y)\,| \le |x|\,|y|$, we obtain the following bound for $D_u f(\xi)$

$$|D_u f(\xi)| \le |\nabla f(\xi)|\,|\mathbf{u}| \;\le |\nabla f(\xi)|$$

as magnitude of unit vector $\mathbf{u}$ is one.

Clearly when $\mathbf{u}$ corresponds to the unit vector obtained by normalizing ∇f (i.e., taking $\frac{\nabla f}{|\nabla f|}$) we see that $|D_u f(\xi)| = |\nabla f(\xi)|$ justifying our stated claim that ∇f gives the magnitude and direction for the maximum rate of change of the function f at a given point.

The above discussion carries forward easily to a function of n independent variables.

$\nabla f(x_1, \ldots, x_n) = (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_2})$

L'Hôpital rule

L'Hôpital rule allows us to evaluate an expression that takes the form $\frac{0}{0}$ or $\frac{\infty}{\infty}$ in the case when the parameter takes a limiting value. L'Hopital rule tells how we can come up with a meaningful value for the above undefined expression by observing the behavior of the function in the vicinity of the limiting value.

Let $f(x)$ and $g(x)$ be given functions of independent variable x. Assume that $f(a) = 0 = g(a)$.

L'Hôpital s rule says that

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(a)}{g'(a)}$$

where $f'(a)$ and $g'(a)$ are the derivatives of the function at the point a.

Example:

$$\lim_{x\to 0} \frac{sinx}{x} = \lim_{x\to 0} \frac{cosx}{1} = 1$$