Author: Timothy A. V. Teatro
Date: Monday 6th February, 2017

PhD Thesis Proposal
EXPLORATION AND ADVANCEMENT OF TECHNIQUES FOR PREDICTIVE CONTROL SYSTEMS SOFTWARE DESIGN AND CONSTRUCTION

1 Introduction

2 A Brief Discussion for the Purpose of Establishing Notation

> We could, of course, use any notation we want; do not laugh at notations; invent them, they are powerful. In fact, mathematics is, to a large extent, invention of better notations.
>
> RICHARD P. FEYNMAN (1918–1988)
> The *great* physicist, my personal hero

UNLESS OTHERWISE indicated, the mathematics in this document is typeset with strict adherence to ISO/IEC 80000-2:2009 (ISO-Math) [1]. Undocumented deviations should be considered a bug. To save the reader the trouble of finding and reading the standard, I shall summarise herein with particular care to point out anything not in common convention in North America.

Quantities which are not variable across time or context (such as immutable constants of mathematics) are set in upright text. For examples, $\exp(1) = \mathrm{e}$ and not the italic e; the ratio of a circle's diameter to its circumference (in flat space) is $\mathrm{\pi}$ and not π; the imaginary unit, defined by $\mathrm{i}^2 = -1$, is not the italic i, and so on. Variables, parameters, contextual constants, running numbers and alike are set in italicized text. For example, $\sum_i a_i = 2 + 4\mathrm{i}$ where i is a counter while i is the imaginary unit.

This rule of italics vs. upright is universally conventional for functions. This is why $\sin(x)$ is correct and the commonly seen italic $sin(x)$ is widely recognized as an error. (Some LaTeX users, particularly those who tend to forget backslashes such as the one in `\sin(x)`, are common offenders.)

Vector and matrix quantities follow the convention of italics vs. upright, and are notated in bold. In addition to being bold, it will generally (but not necessarily) be the case that vectors will be lowercase, while matrices are uppercase:

$$(\boldsymbol{A}\boldsymbol{x})^{\mathsf{T}}\,\boldsymbol{y} = \boldsymbol{x}^{\mathsf{T}}\,(\boldsymbol{A}\boldsymbol{y}). \tag{B.1}$$

(Note, this is the foundational identity of the matrix-transpose of $\boldsymbol{A}$. That is, the matrix $\boldsymbol{A}^{\mathsf{T}}$ that makes the Left Hand Side (LHS) inner product equal to the Right Hand Side (RHS) inner product, for all suitably shaped $\boldsymbol{x}$ and $\boldsymbol{y}$, is by definition the transpose of $\boldsymbol{A}$.)

The $n \times n$ identity matrix is a matrix quantity, (so is set in bold and uppercase), but it is defined for matrices independently of context. As such, its symbol should be bold and upright: $\mathbf{I}_n$[1].

I find the following convention to be the most awkward to North American eyes. I have had reviewers mistake it for poor typesetting despite the great care I take with my documents. *In following with the convention that context independent objects are upright, the differential operators are typeset upright.* For example:

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \partial_x f = \mathrm{e}^{\mathrm{i}x} \Longrightarrow f(x) = -\mathrm{i}\,\mathrm{e}^{\mathrm{i}x} + c,$$

not,

$$\frac{df}{dx} = \partial_x f = \mathrm{e}^{\mathrm{i}x} \Longrightarrow f(x) = -\mathrm{i}\,\mathrm{e}^{\mathrm{i}x} + c.$$

The differences between the upright partial ∂ and *italic* partial ∂ seem trivially subtle. But the upright infinitesimal $\mathrm{d}x$ avoids confusion with the common notation for a distance metric $d(\boldsymbol{x}, \boldsymbol{y})$. The expression $dd/d\boldsymbol{x}$ is nonsense while $\mathrm{d}d/\mathrm{d}\boldsymbol{x}$ is at least meaningful, even if it is distracting.

A good notation should do some work for you. It frees the mind to bring full attention to bare on the real problem, and even carry you part of the way. The legendary mathematician Bertrand Russell is reputed to have said that *a good notation has a subtlety and suggestiveness which at times make it almost seem like a live teacher.*

In the spirit of the previous paragraph, I feel compelled to comment that I dislike this dot: $\boldsymbol{a} \cdot \boldsymbol{b}$ in the context of the vector *dot* product. The great applied mathematician William Gilbert “Gil” Strang even calls it unprofessional [2, p. 108], though I feel that may be overstating it. Worse than useless, that dot distracts from what really is happening. We have been well trained to understand how matrices multiply. Let us simply write the dot product (or inner product) for two column vectors $\boldsymbol{a}^{\mathsf{T}}\boldsymbol{b} = \sum_j a_j b_j$. It is unsurprising then that the outer product (or rank one product, or *tensor* product) is $\boldsymbol{a}\boldsymbol{b}^{\mathsf{T}}$.

The matrix multiplication really lets us connect with the underlying linear algebra. It does work for us, even when we extend linear algebra to functions (that is, functional analysis). Let us think about two functions of a real variable, $f(t)$ and $g(t)$. Let us also discretise by evaluating on a mesh of t that is sampled with interval T. More concisely, $t = kT$ with $k = 0, 1, 2, \ldots$. We can then think of f as a column of values $f_k = f(kT)$. The inner product notation now suggests a convenient form for the inner product of two functions:

$$f^{\mathsf{T}} g = \sum_j f_j\, g_j.$$

[1] This marks a deviation from the ISO/IEC standard which discloses a bold italicised $\boldsymbol{I}$ for this purpose. I view this as an oversight in the standard since it departs from the convention of printing mathematical constants in Roman. I have found good examples where others make this same deviation.

Now, work your way back to the continuous case by allowing $T \longrightarrow 0$, and the above sum condenses to the integral:

$$\lim_{T \longrightarrow 0} \sum_j f_j \, g_j = \int_{-\infty}^{\infty} f(x) \, g(x) \, \mathrm{d}x = \langle f, \, g \rangle.$$

It is short work from here to derive the Fourier or Laplace transformations, and we were well set up for it using the right notation. (If you are unsure about getting Fourier or Laplace out of this, consider taking the inner product of a function of interest with each member of a set of orthogonal basis functions, such as the complex sinusoids. You are decomposing the function vector in a Hilbert space of functions just as one would find the x-, y- and z-components of a vector in Euclidean 3-space.)

I digress for a moment to appreciate that combining the above definition of the inner product with the transpose identity (B.1) can be used to derive integration by parts:

$$(\boldsymbol{A}\,\boldsymbol{x})^{\mathsf{T}}\,\boldsymbol{y} = \boldsymbol{x}^{\mathsf{T}}\,(\boldsymbol{A}\,\boldsymbol{y})$$
$$\int_{-\infty}^{\infty} \frac{\mathrm{d}f}{\mathrm{d}t}\, g(t) \, \mathrm{d}t = \int_{-\infty}^{\infty} f(t) \left(-\frac{\mathrm{d}g}{\mathrm{d}t} \right) \, \mathrm{d}t.$$

The reader can expect two notations for the differential operation. I could hardly be accused of rebellion for adopting the common notation of Leibniz,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\, f,$$

for the derivative of y with respect to x. I will also exercise the more compact notation for the same thing:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \partial_x y.$$

The former notation has the disadvantage of becoming optically obscure if x were to have super/sub-scripts, so in such cases I will fall back to Leibniz.

The notation with the *del*, ∂, is popular in field theory (and related areas of physics) and I believe is based on the notation of Oliver Heaviside which is the same except using a D in place of the del.

The set of real numbers is denoted by $\mathbf{R}$. The symbol $\mathbf{R}_+$ is a common notation for the positive real numbers. It is also a common short hand for the additive group of real numbers, $(\mathbf{R}, +)$. I will surely reference the positive real numbers much more frequently, so the compact notation $\mathbf{R}_+$ is reserved for that. Similarly, $\mathbf{Z}_+$ will represent positive integers, which some may refer to as natural numbers. Still others would not recognize these as natural numbers as there is no general agreement on whether natural numbers should be the positive integers, or the non-negative integers. To avoid confusion, I notate the non-negative integers as $\mathbf{Z}_{\geq 0} = \{0\} \cup \mathbf{Z}_+$, and avoid the phrase *natural numbers* everywhere but this paragraph.

The Kronecker-delta, $\delta^n_m = \delta_{nm}$, (which evaluates to 1 if and only if $n = m$ and 0 otherwise), is upright. So too, is the Dirac-delta, $\delta(x)$, which is zero everywhere except at $x = 0$ with $\int_{-\infty}^{\infty} \delta(x)\, \mathrm{d}x = 1$.

In this document, there will be references to discretely sampled trajectories through (flat) spaces of arbitrary dimension. For example, a trajectory through a 7-dimensional Euclidean space. These sampled trajectories will be collected as a time series into a structure that looks a lot like a matrix: a row of column-vectors forming a rank-2 array. These will typically be notated with a lowercase bold italicized Latin symbol, such as $\boldsymbol{x}$, or $\boldsymbol{u}$—breaking the convention of uppercase for matrices. This departure from the convention indicates that these structures are most usefully thought of as a time series of column-vectors. Take such an entity $\boldsymbol{a} \in \mathbf{R}^{n \times N}$. The

k-th column-vector in the series will be notated as $\boldsymbol{a}^k$ with $0 \le k \le N-1$. Since these are vector quantities, there should be no confusion with exponentiation. The superscript notation frees the subscript position for notational embellishments. This sort of notation is not strange, it is simply borrowed from tensor notation. In tensor notation, there is a difference between super-scripts (which indicate covariant components) and subscripts (which indicate contravariant components). However, we will be working in flat spaces where the two are equivalent. However, I would rather not even confuse the issue with such concerns. Simply keep in mind that superscripts on vectors indicate a column-number in a series of column vectors.

Parentheses, brackets and braces will be used with consistency. Parentheses, $(\,\cdot\,)$, will be used as traditional delimiters.

Braces, $\{\cdot\}$, identify sets. For example, $\{a_1,\ a_2, \dots, a_N\}$. I also make use of the ranged brace notation, which expresses that last set as $\{a_k\}_{k=1}^{N}$.

Brackets, $[\,\cdot\,]$, will be used to enclose composite structures such as vectors and matrices, and they will also be used as outer delimiters for tall set operators, such as $\sum$, $\prod$, $\int$, and $\bigcup$. For example, the definition of the Euclidean metric:

$$d(\boldsymbol{a},\ \boldsymbol{b}) = \|\boldsymbol{a} - \boldsymbol{b}\| = \left(\sum_i \left[a_i - b_i\right]^2\right)^{1/2} \tag{B.2}$$

Of course, that definition for the Euclidean metric is based on the Euclidean, or l^2-norm:

$$\|\boldsymbol{x}\| = \left(\sum_i x_i^2\right)^{1/2}.$$

On a normed space Ω, the norm is written $\|\cdot\|_\Omega$. Similarly, the metric will be written $d_\Omega(\cdot,\ \cdot)$. Of course, that norm and metric may or may not be Euclidean, and the presence of that subscript should incite to question.

The class of n-times continuously differentiable functions is denoted C^n. The derivatives should be bounded so that the supnorm,

$$\|f\|_{\mathrm{C}^k(\Omega)} = \sum_{n=0}^{k} \sup_{x\in\Omega} \left\|\frac{\mathrm{d}^n f}{\mathrm{d}x^n}\right\|,$$

can complete the Banach space, which has important analytical consequences.

2.1 conventions *for* matrix calculus

Two conventions exist for organising the calculus on matrix and vector quantities. Namely, these are the *numerator*- and *denominator*-layouts. The two layouts are related (mostly) by transposition. For this document I adopt the numerator layout, which may be exemplified in the following five cases.

1. For a scalar valued function of a vector, $y : \mathbf{R}^n \longrightarrow \mathbf{R}$, the gradient is a row-vector:
$$\frac{\partial y}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} & \cdots & \frac{\partial y}{\partial x_n} \end{bmatrix}.$$

2. For a vector valued function of a scalar, $\boldsymbol{y} : \mathbf{R} \longrightarrow \mathbf{R}^n$, the derivative is a column-vector:
$$\frac{\partial \boldsymbol{y}}{\partial x} = \begin{bmatrix} \partial y_1/\partial x \\ \partial y_2/\partial x \\ \vdots \\ \partial y_n/\partial x \end{bmatrix}.$$

3. For a vector valued function of a vector, $\boldsymbol{y} : \mathbf{R}^n \longrightarrow \mathbf{R}^m$, the derivative is the matrix:

$$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}.$$

4. For a scalar valued function of a matrix, $y : \mathbf{R}^{m \times n} \longrightarrow \mathbf{R}$, the derivative is a matrix:

$$\frac{\partial y}{\partial \boldsymbol{X}} = \begin{bmatrix} \frac{\partial y}{\partial X_{11}} & \frac{\partial y}{\partial X_{21}} & \cdots & \frac{\partial y}{\partial X_{m1}} \\ \frac{\partial y}{\partial X_{12}} & \frac{\partial y}{\partial X_{22}} & \cdots & \frac{\partial y}{\partial X_{m2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial X_{1n}} & \frac{\partial y}{\partial X_{2n}} & \cdots & \frac{\partial y}{\partial X_{mn}} \end{bmatrix}.$$

5. And finally, for a matrix valued function of a scalar, $\boldsymbol{Y} : \mathbf{R} \longrightarrow \mathbf{R}^{m \times n}$, the derivative (which is only defined in numerator-layout) is a matrix:

$$\frac{\partial \boldsymbol{Y}}{\partial x} = \begin{bmatrix} \frac{\partial Y_{11}}{\partial x} & \frac{\partial Y_{12}}{\partial x} & \cdots & \frac{\partial Y_{1n}}{\partial x} \\ \frac{\partial Y_{21}}{\partial x} & \frac{\partial Y_{22}}{\partial x} & \cdots & \frac{\partial Y_{2n}}{\partial x} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial Y_{m1}}{\partial x} & \frac{\partial Y_{m2}}{\partial x} & \cdots & \frac{\partial Y_{mn}}{\partial x} \end{bmatrix}.$$

A quantity simply stated to be a *vector* should be considered a column-vector unless otherwise indicated.

3 List of Acronyms, Initialisms and Selected Jargon

IEC International Electrotechnical Commission

ISO International Organisation *for* Standardisation

ISO-Math ISO/IEC 80000-2:2009

References

[1] ISO 80000-2:2009. *Quantities and units—Part 2: Mathematical signs and symbols to be used in the natural sciences and technology*. 2009.

[2] W. Gilbert Strang. *Introduction to Linear Algebra*. 4th. Wellesley, MA: Wellesley-Cambridge Press and SIAM, 2009. ISBN: 978-0980232714.

Index