Notes on super-mathematics

Mitchell Porter (mjporter@uq.edu.au)

Supermathematics is the mathematics of supersymmetry, a popular and profound conception in the speculations of theoretical physicists. These notes are a very basic introduction to supermathematics, with a view to machine learning applications.

Based on a presentation to the UQ ML Reading Group, “Applications of super-mathematics to machine learning”

15 March 2019

1 INTRODUCING SUPERNUMBERS

(1.1) There is no real number x with the property $x^2 = -1$.

But we can simply *invent* an “imaginary number” i which is defined by the property $i^2 = -1$.

Then we can combine real and imaginary numbers to make complex numbers a+bi

and a whole new mathematics opens up.

Similarly, all ordinary numbers – real or complex – obey the commutative law

$a\,b = b\,a$

But we can say, let there be special numbers which obey the anticommutative law

$a\,b = -\,b\,a$

Such numbers are called grassmann numbers after their inventor Hermann Grassmann

(who lived in 19th-century Germany)

and are usually written as: θ.

Right away we can see that $\theta\,\theta = -\,\theta\,\theta$, and so $\theta^2 = 0$

and this gives us a new algebra:

$(1 + \theta)\,(1 + \theta) = (1 + \theta) + \theta\,(1 + \theta)$
$= 1 + \theta + \theta + \theta^2$
$= 1 + 2\,\theta$

(1.2) Just as we defined a complex number as the sum of a real number and an imaginary number with a real coefficient,

we may define a supernumber as the sum of a complex number with products of grassmann numbers with complex coefficients.

Usually we want more than one grassmann unit, denoted θ_1, θ_2, θ_3... with the properties

$\theta^2 = 0$

$\theta_i \theta_j = - \theta_j \theta_i$

If we consider e.g. $\theta_1 \theta_2 \theta_1$

$(\theta_1 \theta_2) \theta_1 = (- \theta_2 \theta_1) \theta_1 = - \theta_2 (\theta_1 \theta_1) = - \theta_2 . 0 = 0$

The only products of grassmann units that aren't zero, have no repeated factors; and permuting the factors simply gives the same quantity times some power of -1.

So if there are N distinct grassmann units, $\theta_1 \dots \theta_N$, there will be 2^N components in the supernumbers of that algebra.

$a + b \theta_1 + c \theta_2 + d \theta_3$
$+ e \theta_1 \theta_2 + f \theta_1 \theta_3 + g \theta_2 \theta_3 + h \theta_1 \theta_2 \theta_3$

(1.3) Complex numbers give rise to useful new mathematics like complex analysis and complex geometry.

Similarly, supernumbers give rise to superalgebra, superanalysis, supergeometry...

All this was developed primarily for its use in physics.

2 SUPERNUMBERS IN PHYSICS

(2.1) In physics, there are two kinds of particles, bosons and fermions.

Force particles (like the photon) are bosons. The same quantum state can be occupied by more than one boson at the same time. This happens in lasers.

Matter particles (like the electron) are fermions. If a quantum state is already occupied by a fermion, no other fermion can join it. This happens with electron orbitals in atoms.

One way to obtain quantum probabilities is through a 'path integral' which gives a probability for going from A to B, by summing over all possible 'paths' from A to B.

The path integral for bosons involves ordinary numbers, but the path integral for fermions requires grassmann numbers. This was figured out by CANDLIN in 1956.

(2.2) Particle physics is full of symmetries, in which particles are grouped together as partners under some transformation. Eventually the possibility of bosons and fermions in the same symmetry group was discovered and called supersymmetry.

The mathematics of supernumbers was developed by BEREZIN, to describe supersymmetry.

Experiment has failed to reveal superpartners of the known particles.

However, supermathematics has found use outside particle physics, which I will divide into algebraic and geometric uses.

3 ALGEBRAIC USE OF SUPERMATH: STOCHASTIC SYSTEMS

Along with using grassmann numbers to describe actual fermions,

a method was employed in physical calculation, whereby fictitious fermions called 'ghosts' were introduced, to compensate for the effects of 'gauge-fixing', an artificial symmetry-breaking also done just for the purposes of calculation.

The 'ghost fermions' introduced an overall factor of '1' to the path integral, so they changed nothing about the final predictions, but they cancelled out side effects of the gauge-fixing that needed to be cancelled out.

It was later discovered by PARISI and SOURLAS that grassmann variables could similarly be added to certain stochastic differential equations, creating supersymmetric equations.

This application is reviewed by OVCHINNIKOV.

I call this an 'algebraic' application of supermathematics, because it uses the cancellations that are common in supersymmetric systems.

4 GEOMETRIC USE OF SUPERMATH: DIFFERENTIAL GEOMETRY AND MORSE THEORY

(4.1) Another place where grassmann algebra turns up is in differential geometry.

Three-dimensional vector algebra already contains a 'cross product' which is anticommutative:

a x b = - b x a

This is because the cross product really defines an 'oriented parallelogram', whose sides are the vectors a and b, and whose 'orientation' is another vector perpendicular to the parallelogram.

(More precisely, that parallelogram is the 'exterior product' of two vectors, written a ^ b,

and the usual cross product is the 'Hodge dual' of the exterior product.)

(4.2) The exterior product allows us to build up oriented cell-like objects of higher and higher dimension.

The exterior product of two vectors a ^ b lies in a plane,

the exterior product of three vectors a ^ b ^ c lies in a three-dimensional hyperplane...

and in fact, when combined by the exterior product, the unit vectors behave like the grassmann units θ_1, θ_2... introduced earlier.

These multivector objects have various uses in geometry. In particular, they provide a coordinate-independent way to describe tensor calculus.

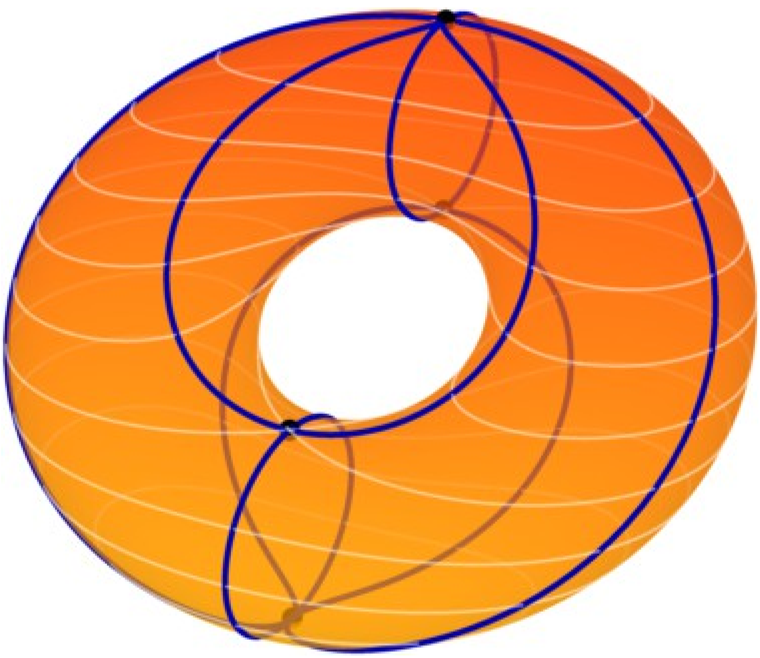
In this case, one takes the outer product of infinitesimal vectors, obtaining e.g. dx ^ dy.

(4.3) The geometer Marston Morse developed a kind of algebraic topology, now called Morse theory,

in which a skeletal approximation of a manifold (below, a torus) is obtained by defining a smooth function everywhere on the manifold,

and then considering turning points of the function ('critical points'), and lines of gradient flow connecting the critical points.

The 'index' of a critical point is the number of directions in which Morse's function decreases at that point. The indexes of the critical points tell us how many linearly independent forms exist on the manifold, and this in turn tells us its topology – how many holes, and what kind of holes, it contains.



The physicist Edward Witten discovered that, for a given manifold and Morse function, he could define a supersymmetric quantum system which encoded the same information.

The even-dimensional forms were bosons, the odd-dimensional forms were fermions; the critical points corresponded to physical ground states; and the lines of flow between the critical points, corresponded to quantum tunneling between the ground states.

5 USE OF SUPERMATH IN MACHINE LEARNING?

(5.1) These are concepts and analytical tools which may facilitate design of new ML systems. (M. Kamgarpour suggests the name 'super data science' for this hypothetical new subdiscipline.)

The actual idea that supersymmetry might be applied to machine learning, is due to Jordan Bennett of the University of the West Indies.

He has so far unearthed two examples of supersymmetric techniques being employed in the study of ML systems.

… YAMAZAKI 2000 (Hopfield networks)

… ALGORITSAS ET AL 2017 ('random perceptrons')

In both cases, it is an 'algebraic' application of supersymmetry, to prove something about a stochastic ML system.

Bennett himself makes a variety of suggestions, including:

… Learn supersymmetric weights

… Learn a submanifold of a supermanifold

… Supersymmetrize the Boltzmann machine

(5.2) There is also an ML paradigm known as 'persistent homology machine learning' (PHML).

Persistent homology studies the 'persistence' of topological features as the description of an object becomes more and more coarse-grained. It is considered to mediate between topology and geometry.

Morse theory has already been applied in PHML, so there is a good chance that PHML can use the supersymmetric approach to Morse theory.

(5.3) I also consider the work of Igor Ovchinnikov, already mentioned as an 'algebraic' application of supersymmetry, to be very promising here. In Ovchinnikov's hands, supersymmetry provides a general theory of dynamical systems, offering new perspectives on phenomena such as deterministic chaos and 1/f noise. In a sense, he unifies the algebraic and geometric applications of supersymmetry, since he interprets stochastic supersymmetry in terms of Witten's topological supersymmetry.

References

Algoritsas et al, “Out-of-equilibrium dynamical mean-field equations for the perceptron model”
https://arxiv.org/abs/1710.04894

Bennett, “Thought Curvature”
https://openreview.net/pdf?id=S1aotAc8M

Candlin, “On sums over trajectories for systems with Fermi statistics”
http://dx.doi.org/10.1007/BF02745446

Karabegov et al,“Felix Alexandrovich Berezin and his work”
https://arxiv.org/abs/1202.3930

Ovchinnikov, “Introduction to Supersymmetric Theory of Stochastics”
https://arxiv.org/abs/1511.03393

Parisi & Sourlas, “Supersymmetric field theories and stochastic differential equations”
https://doi.org/10.1016/0550-3213(82)90538-7

Witten, “Supersymmetry and Morse Theory”
https://projecteuclid.org/download/pdf_1/euclid.jdg/1214437492

Yamazaki, “Information dynamics of neural networks with the aid of supersymmetry fields as themicroscopic thermal flow”
https://doi.org/10.1088/0305-4470/33/47/306

Sketch of the Morse complex for a torus from Wikipedia
https://en.wikipedia.org/wiki/File:TiltedTorusFlowLines.png