Gradiency: A Two-Tier Introduction

By Arthur Shevenyonov

Abstract

The positivist programme of data mining may usher in rather unstable results depending on how observability is defined—or whether strictly sequential (and hence largely orthogonal) hypothesis construction-and-testing can meaningfully be ensured. This could be due to a flawed approach to refutation (a special if naïve case of inverted syllogistic implication) and OLS or GLS type modeling (with factors presence alone or their levels representation not quite doing the job of robust and observation-laden functional stretching or local approximation in small differences). One other reason is the tossup over inferential significance versus power, as one of the numerous related tradeoffs to be overcome. A dichotomy of Gradiency and Azimuthality is proposed, which points to an essential duality reduction that can be extended all the way up to securing a paradigm conducive to ontological convergence.

Background & Motivation

What could be the ultimate measure of modeling adequacy? Could a choice of paradigm render this choice of second-order importance or actually drive it? The former issue depends on what one means by or expects of a model, in line with the prevailing [static or irrefutable] paradigm or indeed the [exogenously imposed] scientific method—which in turn is somehow prone to a pragmatic or instrumentalist bias from the outset.

When embarking on a positivist approach—with applicable metrics being as diverse as, goodness of fit (striking a balance of alpha significance versus beta power), Cronbach's alpha (referring to the invariance or validity of categories being used), or iterative calibration—it is up to prior scrutiny of whether one of these could suffice or all should be treated as a necessary cut-off. Not only does strong-form rejection of an omnibus null yield but weak inference (i.e. refuting one of the coefficients rather than all of them, with a conjunction operator implied *ex ante*), its failure to arrive at a *specific* non-zero coefficient value further adds to the inconclusive syllogistic exercises: If A means B, and "not B" means "not A," then what could "non not B" tell us of A's existential mode?

The conventional positive routine would be split into two, presumably disjoint sections:

$$\emptyset \to H$$

$$H \to M \equiv \cap \{M_i\}$$

Whereas the former part pertains to hypothesis formation (perhaps valid insofar as is the respective syllogism of the form, $\emptyset \rightarrow 1$), the latter amounts to a testable set of implications, presumably related by an "AND" operator as a kind of quasi-group (drawing upon no unity or inverse other than refutation). In a setup such as this one, refuting M amounts to a weak condition, i.e. questioning one of the respective implications—which makes it even more questionable to discard the original hypothesis on the weakness of its minor detail alone. In fact, this may work in natural sciences, though only insofar as the implication set can be seen as complete—which it never is at any particular point in time. One way of getting around the issue would be to extend the programme to a circular, multi-stage setup, with H being *updated* rather than rejected on minor refutation. It should come as little surprise if such an atheoretical stance on hypothesis formation only begs further questions.

Worse yet, it is unclear how to approach the issue whenever the validity of the very mapping in question is sub-unity, e.g. probabilistic in nature? In fact, one need not refer to fuzzy or quantum logic of any sort to question the very criterion of *transferability* as one pillar of the model's inner *consistency*—as long as syllogistic implication can be extended. For instance, A maps into B, B into C and so forth up through G might suggest A maps into G, or that refuting G amounts to questioning A. However, though holding *locally* (as between adjacent points such as B and C), this inference is problematic for mutually *distant* categories (e.g. A and C or B and G). Furthermore, as long as a probabilistic merit can be assigned to any such local transfer, it would be awkward *technically* to speculate of transferability, if only in light of the resultant product of partial probabilities serving as a rough proxy for the likelihood of the implied composite event (subject to joint or conditional distribution or Bayesian updates), clearly falling far too short of unity:

$$A \overset{p}{\rightarrow} B \overset{q}{\rightarrow} C \ldots \overset{r}{\rightarrow} G$$

$$P\{A \rightarrow G\} \sim \prod_i P\{X_i \rightarrow X_{i+1}\} = p * q * \ldots * r \ll 1$$

In a well-defined sense, insofar as A could be seen as a rethinking of H, the interim set of *(B..F)* could act to generalize the aforementioned set of implications M. At this rate, mapping into G as an ultimate reality further generalizes the refinements of H.

The very *(A,G) path* may or may not prove to be *effective* from the standpoint of the resultant probabilistic-product validity. What is more, though, its *interim trajectory*—referring to the choice of those inner categories or indeed conceptual "azimuths"—may not turn out to be *efficient* with an eye on the number of those ancillary roadblocks opted for. Now, this number or "power" is related to the scale of inference (or how far it extends and how questionable a refuting of its distant layers would be), whereas the local or transitional odds can only be coupled with it indirectly.

It remains to be seen whether this duality or efficacy-efficiency tradeoff resembles that of alpha versus beta (or Type I versus Type II error). For now, suffice it to surmise that a similar transferability (or at least *commutativity*) loss could plague any differential calculus, as long as its linearizing symmetry only works locally, i.e. collapses for non-infinitesimal deferents.

Azimuthality versus Gradiency

If one were to abide by the above schema, it would be straightforward to favor an efficient as well as effective *(A, G)* path—be it a mechanism or modality—of attaining a reasonably accurate depiction of reality G by starting with a prior metaphor of G, or model A. This is what I have referred to as Gradiency from way back in 1999, which may either point to G as an ultimate model (if not *the* reality) or to *(A, G)* as an optimal path of attaining one—or at any rate, of gliding along the "surface" of reality such that a *systemic* error (any delta G) would be preferred as a second-best over "noise" (standing for either stochasticity or ad-hoc deference) as long as the path is attached with a parsimonious algorithm for recovering the original or undoing the gap. In a sense, this amounts to the likes of a tensor, a functional or operator inverse, or syllogistic inversion—even though all of these proxies are largely helpless when capturing the generalized notion of Gradiency, and have been made no use of.

Originally, the latter would resemble concepts as seemingly diverse as tangency in economics (e.g. equimarginal conditions of optimization, akin to Newtonian conservation laws), weak non-Euclidean parallelism, or even the classic *mimesis*. In the latter case, it is straightforward to appreciate how those B through F could either serve as [variable candidate] instruments (as opposed to a method all set in stone) or interim *eidoi* (in Platonic lingo)—whose relevance only hinges on whether or not the ultimate path could be posited akin to Feynman's path interval or in terms of path-invariance as in functional analysis.

In fact, Gradiency is largely about the success of picking the initial A as the representative guess which is supposed to lead to G as fast (i.e. with as few interim variables) as possible. This suggests rendering the design prone to minimal *(A, G)* distance or otherwise fast convergence—which simultaneously facilitates "distant" refutation.

On second thought, there is not need to dwell on either the positivist apparatus or those individual proxies for the new and complete category; much less because neither has underpinned the latter in retrospect. For one thing, *Gradiency* could roughly be thought of as either maximal performance on *both* ends (efficiency and efficacy) or otherwise an *optimal* trade-off between the two—with an eye on tapping the ontology and not empirical phenomenology per se. One should bear in mind that, apart from the criteria of technical observability, their shift would amount to an error or shock which is not necessarily of an additive or linear nature, or at any rate need not amount to a systematic gap that could be undone or sorted out at a low cost.

For instance, a thought experiment readily suggests how an implied data pattern could reveal a very different functional form or, for that matter, extra dimensions as well as added inter-linkage (main versus interactive effects). Whereas the ordinary least squares (OLS) regression or its GLS extensions are not impeccable conceptually (given that a process can hardly be approximated globally or in levels terms the way it can be rendered so locally or in small differences), the very presence of key dimensions (however complete) does not outweigh the importance of their actual accommodating functional form—which may vary drastically as the observability frontier shifts.

By contrast, *Azimuthality* could be treated in terms of either low efficiency (too many interim building blocks) or mediocre efficacy (slow or uncertain convergence to the true "surface" or structure), or both—referring to a suboptimal tradeoff between the two. Needless to say, this implied *duality* between the two is but a partial or one-faceted reduction, with either one capturing a far more full-fledged set of representations as well as implications than that overlapping subset as allowed by direct contrasting. Nor does the Azimuthality-Gradiency dichotomy quite come down to the *(A,G)* distance or juxtaposition, even though there might well be some rationale behind this proposition—again marking the duality of attempting an initial (or distant, suboptimal) representation of reality (A) versus attaining its actual frontier or close match (G).

In a sense, Azimuthality may not refer to deviation per se, but rather to a path—or indeed paradigm—that tends to either overlook ontology altogether (the *why* question not least) or sweep under the rug any excessive complexity of bridging the gap, while positing this inefficiency as "riches" and attainment in its own right. For instance, the mainstream algebraic paradigm works in a peculiar yet long-accustomed-to way, whereby the less important or seemingly straightforward results turn out to be the hardest to demonstrate. This does not merely point to a level of rigor being overshot, yet it does speak volumes to the excessive cost in terms of interim results being maintained as an "achievement in its own right," while essentially amounting to that interior or core of *(B, F)* as one other distance representation, which could alternatively be posited as path dependence of the sort:

$$X_i = A(X_{i-1})$$

By reducing this to a recursive or functional solution and applying the corner conditions as well as the "*ab nihilo*" (positivism-compliant approach whereby the source of the initial hypothesis is either unknown or irrelevant without questioning the syllogistic validity), it follows that:

$$(\emptyset \to A) = True = P_0 \ or \ (\emptyset \overset{p_0}{\to} A) \sim P\{A_0, A\}$$

$$X_i = A^{[i]}(X_0) \equiv A^{[i]}(\emptyset) \equiv A^{[i-1]}(A) \sim \emptyset * \prod_{k=0}^{i} (X_k, X_{k+1})$$

$$\exists m: G = X_m = A^m(\emptyset) = A^{m-1}(A) \sim \emptyset * (\emptyset, A) * (A, B) * \ldots * (F, G)$$

In passing, one might have liked to postulate $G \to A$ as the prior stage in place of $\emptyset \to A$, though that would be another stage of analysis, indeed a major chunk of future research in its own right. In some latent sense, as will be shown in forthcoming announcements, this effective $G \to G$ programme for Gradiency might lend itself with Schelling's (1801) "identity" philosophy (Vater, 2001, pp. 339-371). For better or worse, in no manner has this latter strand of literature (which I was referred to days ago) affected the proposed approach, and nor will further related research build on Gradiency alone.

One way or another, the choice of each interim building block has been affected by the grand enterprise or design A whose support or refutation hinges on G match yet not any interim goodness of fit. In other words, the closest-to-G interim implications or categories are exactly those easiest to refute yet least relevant to refuting A. By the same token, the validities of local transitions from X_i to X_{i+1}, or their prior versus posterior odds gap, are of far lesser relevance than the resultant composite validity—or indeed its posterior unity versus zero, knife-edge probabilistic equivalent of attaining G versus failure to do so.

Whereas Gradiency is more in line with there being a “most elegant and fruitful path” (indeed a facilitating paradigm) as chosen from the outset, Azimuthality may hover anywhere around incessant deferents, ad-hoc calibration, or otherwise satisficing that may or may not be at odds with path dependence (e.g. the paradigm’s internal consistency as one notion of parsimony). I would let the jaded reader decide whichever metaphor is the closer take on Azimuthality: A karate-do master doing the arcane *kata* exercise or a Standard Model theorist feeling happy about arriving at another set of fudge dimensions with which to back a 5% explanatory power on “non-dark” matter and energy.

Future Research

The proposed dichotomy of Azimuthality versus Gradiency will be shown to provide a special account of a far broader domain of calculus that might mark a paradigm shift toward greater inferential facility. It may take a rethinking of the positivist premises (which have thus far been perceived to act as a simplifying bridge), if one is to arrive at cross-disciplinary solutions that show promise of overcoming tradeoffs of the Heisenberg type, or between significance versus power (Type I and II errors), or between efficiency versus efficacy.

By far the only serious cost to be considered is the questionable and myopic savings accruing from a discarding of ontology, or a frontier of generalization acting as the efficacy or explanatory slack. Future research will dwell on the notion of Gradiency in greater detail, while naturally motivating and informing how its linkage with supplementary (yet more ultimate) categories suggests that, *too costly means of lesser relevance*—which has little to do with either instrumental pragmatics or probabilistic interpretations.

For now, it is important to appreciate how Gradiency pertains to “structural co-movement” as preset by a choice of paradigm, rather than ad-hoc or model-specific vector collinearity, statistical correlation, weak topological invariance, or group-theoretic fit to name but a few.

References

Schelling, F. (2001 [1801]). Darstellung des Systems Meiner Philosophie. (M. Vater, Trans.). *Philosophical Forum*, *32*(4), 339-371.