

From analyzing data to studying observers

A journey through four recent papers

Patrizio Frosini

Department of Mathematics, ALMA-AI, AM² and ARCES, University of Bologna
patrizio.frosini@unibo.it

“Geometry of Deep Learning” Ph.D. Course,
CaLIGOLA Training School, FIM, Modena, February 23, 2023

Outline

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOS

Building linear and nonlinear GENEOS

How can we use GENEOS in applications?

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOS

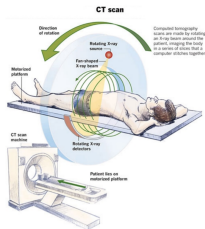
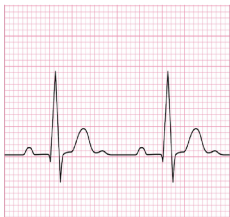
Building linear and nonlinear GENEOS

How can we use GENEOS in applications?

Data can be often regarded as functions

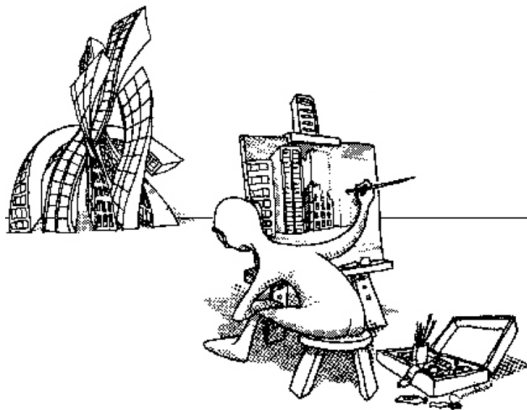
Some examples of data that can be seen as functions:

- An electrocardiogram (a function from \mathbb{R} to \mathbb{R});
- A gray-level image (a function from \mathbb{R}^2 to \mathbb{R});
- A computerized tomography (CT) scan (a function from a helix to \mathbb{R}).



Data are processed by observers

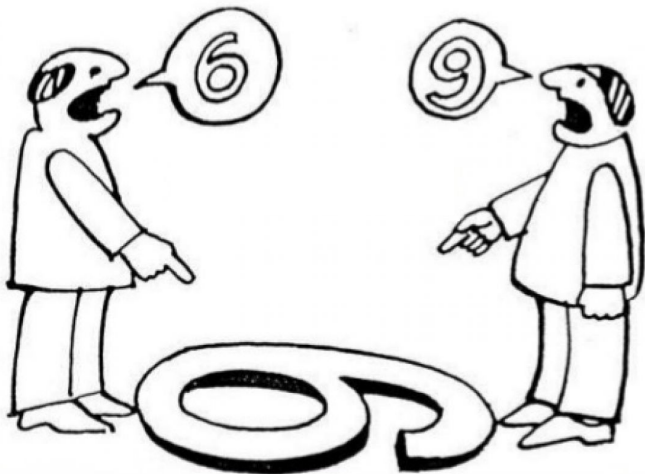
Data have no meaning if no observer elaborates them.



An observer is an agent that transforms data.

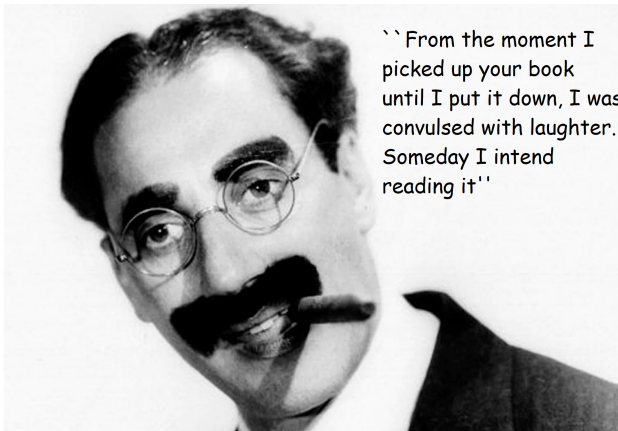
Observers are variables in data analysis

Data interpretation strongly depends on the chosen observer:



Our interest in data is greatly overrated

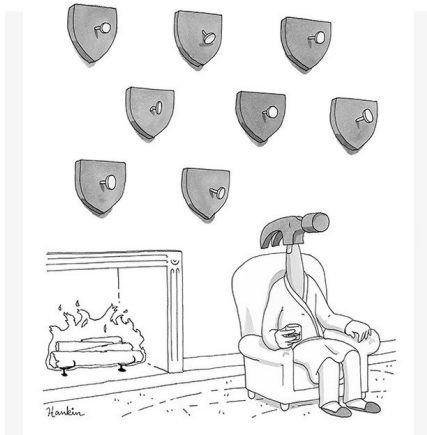
We are hardly ever interested directly in data but in the reaction of the observer to the presence of data.



`` From the moment I
picked up your book
until I put it down, I was
convulsed with laughter.
Someday I intend
reading it''

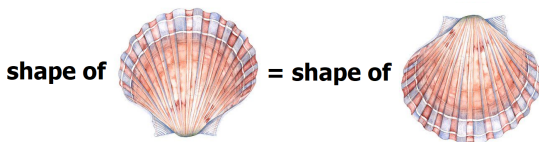
No data structure

Generally speaking, there is no structure in data. The structure of data is a projection of the structure of the observer.

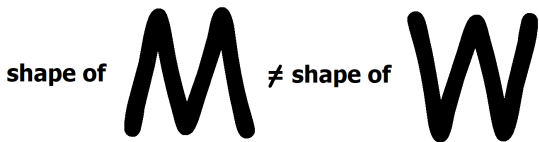


Observers are often associated with invariance groups

Observers often think that some data are equivalent to each other, according to an invariance group.



The group G is not established once and forever: when the observer changes, G changes too.



Representing observers as equivariant operators

Observers are structures able to change data into other data, and usually do that by respecting some data equivalences, i.e., by commuting with some transformations.

As a first approximation, observers can be represented as group equivariant operators (GEOs).

In this talk we will illustrate some results on the theory of **Group Equivariant Non-Expansive Operators (GENEOs)**.

Why “non-expansive”?

Because

1. observers are often assumed to simplify the metric structure of data in order to produce meaningful interpretations;
2. non-expansiveness guarantees good topological properties.

The key role of observers in data analysis





Topological and metric basics for the theory of GENEOS

Building linear and nonlinear GENEOS

How can we use GENEOS in applications?

How could we represent observers?

Towards a topological-geometrical theory of group equivariant non-expansive operators for data analysis and machine learning

Mattia G. Bergomi ¹, Patrizio Frosini ^{2,3*}, Daniela Giorgi ⁴ and Nicola Quercioli ^{2,3}

We provide a general mathematical framework for group and set equivariance in machine learning. We define group equivariant non-expansive operators (GENEOs) as maps between function spaces associated with groups of transformations. We study the topological and metric properties of the space of GENEOs to evaluate their approximating power and set the basis for general strategies to initialize and compose operators. We define suitable pseudo-metrics for the function spaces, the equivariance groups and the set of non-expansive operators. We prove that, under suitable assumptions, the space of GENEOs is compact and convex. These results provide fundamental guarantees in a machine learning perspective. By considering isometry-equivariant non-expansive operators, we describe a simple strategy to select and sample operators. Thereafter, we show how selected and sampled operators can be used both to perform classical metric learning and to inject knowledge in artificial neural networks.

<https://rdcu.be/bP6HV>

All begins with the space of admissible functions

Let X be a nonempty set. Let Φ be a topological subspace of the set \mathbb{R}_b^X of all bounded functions φ from X to \mathbb{R} , endowed with the topology induced by the metric

$$D_\Phi(\varphi_1, \varphi_2) := \|\varphi_1 - \varphi_2\|_\infty.$$

We can see X as the space where we can make our measurements, and Φ as the space of all possible measurements. We will say that Φ is the *set of admissible functions*. In other words, Φ **is the set of all functions from X to \mathbb{R} that can be produced by our measuring instruments (or by other observers)**. For example, a gray-level image can be represented as a function from the real plane to the interval $[0, 1]$ (in this case $X = \mathbb{R}^2$).

Perception pairs

Let us consider a group G of bijections $g : X \rightarrow X$ such that $\varphi \in \Phi \implies \varphi \circ g \in \Phi$ for every $\varphi \in \Phi$. We say that (Φ, G) is a **perception pair**.

The choice of a perception pair states which data can be considered as legitimate measurements (the functions in Φ) and which group represents the equivalence between data (the group G).

To proceed, we need to introduce suitable topologies on X and G . Before doing that, we recall that the **initial topology** τ_{in} on X with respect to Φ is the coarsest topology on X such that every function φ in Φ is continuous.

A pseudo-metric on X

Let us define on X the pseudo-metric

$$D_X(x_1, x_2) = \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)|.$$

D_X induces a topology τ_{D_X} on X .

Theorem

The topology τ_{D_X} is finer than the initial topology τ_{in} on X with respect to Φ . If Φ is totally bounded, then τ_{D_X} coincides with τ_{in} .

The use of D_X implies that we can distinguish two points only if a measurement exists, taking those points to different values.

A pseudo-metric on X

The following properties are of use in our model.

Theorem

Every function in Φ is non-expansive, and hence continuous.

Theorem

If Φ is compact and X is complete, then X is compact.

In the following, we will usually assume that Φ is compact and X is complete (and hence compact).

Some magic happens: each bijection is an isometry

Let $\text{Bij}(X)$ be the group of all bijections from X to X , and denote by $\text{Bij}_\Phi(X)$ the subgroup of all $g \in \text{Bij}(X)$ such that $\varphi \circ g \in \Phi$ and $\varphi \circ g^{-1} \in \Phi$ for every $\varphi \in \Phi$. Let $\text{Homeo}(X)$ be the group of all homeomorphisms from X to X with respect to D_X , and denote by $\text{Homeo}_\Phi(X)$ the subgroup of all $g \in \text{Homeo}(X)$ such that $\varphi \circ g \in \Phi$ and $\varphi \circ g^{-1} \in \Phi$ for every $\varphi \in \Phi$. Let $\text{Iso}(X)$ be the group of all isometries from X to X , and denote by $\text{Iso}_\Phi(X)$ the subgroup of all $g \in \text{Iso}(X)$ such that $\varphi \circ g \in \Phi$ and $\varphi \circ g^{-1} \in \Phi$ for every $\varphi \in \Phi$.

Proposition

$$\text{Bij}_\Phi(X) = \text{Homeo}_\Phi(X) = \text{Iso}_\Phi(X).$$

A pseudo-metric on G

Let us now focus our attention on a subgroup G of $\text{Homeo}_\Phi(X)$.

We can define a pseudo-metric D_G on G by setting

$$D_G(g_1, g_2) := \sup_{\varphi \in \Phi} D_\Phi(\varphi \circ g_1, \varphi \circ g_2).$$

Theorem

G is a topological group with respect to D_G and the action of G on Φ by right composition is continuous.

Theorem

If Φ is compact and G is complete, then G is compact.

GEOs and GENEOS

Each pair (Φ, G) with $G \subseteq \text{Homeo}_\Phi(X)$ is called a *perception pair*.

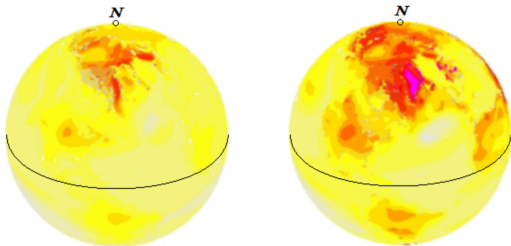
Let us assume that two perception pairs (Φ, G) , (Ψ, H) are given, and fix a group homomorphism $T : G \rightarrow H$.

Each function $F : \Phi \rightarrow \Psi$ such that $F(\varphi \circ g) = F(\varphi) \circ T(g)$ for every $\varphi \in \Phi, g \in G$ is called a *Group Equivariant Operator (GEO)* associated with the homomorphism T .

If F is also non-expansive (i.e., $D_\Psi(F(\varphi_1), F(\varphi_2)) \leq D_\Phi(\varphi_1, \varphi_2)$ for every $\varphi_1, \varphi_2 \in \Phi$), then F is called a *Group Equivariant Non-Expansive Operator (GENEO)* associated with the homomorphism T .

An example of GENE0

Let us assume to be interested in the comparison of the **distributions of temperatures** on a sphere, taken at two different times:



Let us also assume that only two opposite points N, S can be localized on the sphere.

An example of GENEIO

Let us introduce two perception pairs $(\Phi, G), (\Psi, H)$ by setting

- $X = S^2$
- $\Phi =$ set of 1-Lipschitz functions from S^2 to a fixed interval $[a, b]$
- $G =$ group of rotations of S^2 around the axis $N - S$

and

- $Y =$ the equator S^1 of S^2
- $\Psi =$ set of 1-Lipschitz functions from S^1 to $[a, b]$
- $H =$ group of rotations of S^1

An example of GENEIO

This is a simple example of GENEIO from (Φ, G) to (Ψ, H) :

- $T(g)$ is the rotation $h \in H$ of the equator S^1 that is induced by the rotation g of S^2 , for every $g \in G$.
- $F(\varphi)$ is the function ψ that takes each point y belonging to the equator S^1 to the average of the temperatures along the meridian containing y , for every $\varphi \in \Phi$;

We can easily check that F verifies the properties defining the concept of group equivariant non-expansive operator with respect to the isomorphism $T : G \rightarrow H$.

In plain words, our GENEIO simplifies the data by transforming “temperature distributions on the earth” into “temperature distributions on the equator”.

Two key results (and two good news for applications)

Let us assume that a homomorphism $T : G \rightarrow H$ has been fixed.
Let us define a metric D_{GENEO} on $\text{GENEO}((\Phi, G), (\Psi, H))$ by setting

$$D_{\text{GENEO}}(F_1, F_2) := \sup_{\varphi \in \Phi} D_{\Psi}(F_1(\varphi), F_2(\varphi)).$$

Theorem

If Φ and Ψ are compact, then $\text{GENEO}((\Phi, G), (\Psi, H))$ is compact with respect to D_{GENEO} .

Theorem

If Ψ is convex, then $\text{GENEO}((\Phi, G), (\Psi, H))$ is convex.

Two key observations (1)

- While the space of data is often non-convex (and hence averaging data does not make sense), the assumption of convexity of Ψ implies the convexity of the space of observers and allows us to consider the “average of observers”.



Two key observations (2)

- Our main goal is to develop a good geometric and compositional theory to approximate an ideal observer. In our model, "approximate an observer" means to look for a GENE F that minimizes a suitable "cost function" $c(F)$. The cost function quantifies the error that is committed by taking the GENE F instead of the ideal observer. Since the space of GENE F is compact and convex (under the assumption that the data spaces are compact and convex), if the cost function $c(F)$ is strictly convex we have that there is one and only one GENE that best approximates the ideal observer.

The key role of observers in data analysis

Topological and metric basics for the theory of GENEOS

Building linear and nonlinear GENEOS

How can we use GENEOS in applications?

How can we build linear and nonlinear GENEOSs?



On the Construction of Group Equivariant Non-Expansive Operators *via* Permutants and Symmetric Functions

Francesco Conti^{1,2}, Patrizio Frosini^{3,4,5,6*} and Nicola Quercioli^{3,7}

¹ Department of Mathematics, University of Pisa, Pisa, Italy, ² Institute of Information Science and Technologies "A. Faedo", National Research Council of Italy (CNR), Pisa, Italy, ³ Department of Mathematics, University of Bologna, Bologna, Italy, ⁴ Alma Mater Research Center on Applied Mathematics, University of Bologna, Bologna, Italy, ⁵ Alma Mater Research Institute for Human-Centered Artificial Intelligence, University of Bologna, Bologna, Italy, ⁶ Research Centre on Electronic Systems for the Information and Communication Technology, University of Bologna, Bologna, Italy, ⁷ ENEA Centro Ricerche Bologna, Bologna, Italy

<https://www.frontiersin.org/articles/10.3389/frai.2022.786091/full>

Elementary methods to build GNEOs

Proposition (Composition)

If $F_1 \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. $T_1 : G \rightarrow H$ and $F_2 \in \text{GENEO}((\Psi, H), (\chi, K))$ w.r.t. $T_2 : H \rightarrow K$ then $F_2 \circ F_1 \in \text{GENEO}((\Phi, G), (\chi, K))$ w.r.t. $T_2 \circ T_1 : G \rightarrow K$.

Proposition (Image by a 1-Lipschitz function)

If $F_1, \dots, F_n \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. $T : G \rightarrow H$, L is a 1-Lipschitz map from \mathbb{R}^n to \mathbb{R} , and $L^(F_1, \dots, F_n)(\Phi) \subseteq \Psi$ (where L^* is the map induced by L), then $L^*(F_1, \dots, F_n) \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. T .*

The next three statements follow from the last proposition.

Elementary methods to build GENEOS

Proposition (LATTICE OF GENEOS)

If $F_1, \dots, F_n \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. $T : G \rightarrow H$ and $\max(F_1, \dots, F_n)(\Phi), \min(F_1, \dots, F_n)(\Phi) \subseteq \Psi$, then $\max(F_1, \dots, F_n), \min(F_1, \dots, F_n) \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. T .

Proposition (Translation)

If $F \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. $T : G \rightarrow H$, and $F_b(\Phi) \subseteq \Psi$ for $F_b(\varphi) := F(\varphi) - b$, then $F_b \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. T .

Proposition (Convex combination)

If $F_1, \dots, F_n \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. $T : G \rightarrow H$, $(a_1, \dots, a_n) \in \mathbb{R}^n$ con $\sum_{i=1}^n |a_i| \leq 1$ and $F_\Sigma(\Phi) \subseteq \Psi$ for $F_\Sigma(\varphi) := \sum_{i=1}^n a_i F_i(\varphi)$, then $F_\Sigma \in \text{GENEO}((\Phi, G), (\Psi, H))$ w.r.t. T .

Permutant measures

Let us consider the set $\Phi = \mathbb{R}^X \cong \mathbb{R}^n$ of all functions from a finite set $X = \{x_1, \dots, x_n\}$ to \mathbb{R} , and a subgroup G of the group $\text{Bij}(X)$ of all permutations of X .

Definition

A finite (signed) measure μ on $\text{Bij}(X)$ is called a *permutant measure* with respect to G if every subset H of $\text{Bij}(X)$ is measurable and μ is invariant under the conjugation action of G (i.e., $\mu(H) = \mu(gHg^{-1})$ for every $g \in G$).

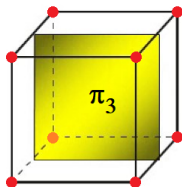
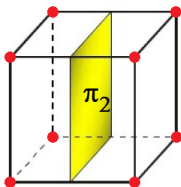
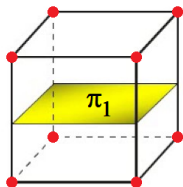
Proposition

If μ is a permutant measure with respect to G , then the map $F_\mu : \mathbb{R}^X \rightarrow \mathbb{R}^X$ defined by setting $F_\mu(\varphi) := \sum_{h \in \text{Bij}(X)} \varphi h^{-1} \mu(h)$ is a linear GEO. If $\sum_{h \in \text{Bij}(X)} |\mu(h)| \leq 1$, then $F_\mu(\varphi)$ is a GENEQ.

An example of permutant measure

Let us consider the set X of the vertices of a cube in \mathbb{R}^3 , and the group G of the orientation-preserving isometries of \mathbb{R}^3 that take X to X . Let π_1, π_2, π_3 be the three planes that contain the center of mass of X and are parallel to a face of the cube. Let $h_i : X \rightarrow X$ be the orthogonal symmetry with respect to π_i , for $i \in \{1, 2, 3\}$.

We can now define a permutant measure μ on the group $\text{Bij}(X)$ by setting $\mu(h_1) = \mu(h_2) = \mu(h_3) = c$, where c is a positive real number, and $\mu(h) = 0$ for any $h \in \text{Bij}(X)$ with $h \notin \{h_1, h_2, h_3\}$.



Building GENEOS by permutant measures

It is interesting to observe that the set $\text{PM}(G)$ of permutant measures with respect to G is a lattice. Indeed, if $\mu_1, \mu_2 \in \text{PM}(G)$, then the measures μ', μ'' on $\text{Bij}(X)$, respectively defined by setting $\mu'(h) := \min\{\mu_1(h), \mu_2(h)\}$ and $\mu''(h) := \max\{\mu_1(h), \mu_2(h)\}$, still belong to $\text{PM}(G)$. Moreover, if $\mu \in \text{PM}(G)$ then $|\mu| \in \text{PM}(G)$. Furthermore, $\text{PM}(G)$ is closed under linear combination. Therefore, $\text{PM}(G)$ has a natural structure of real vector space.

We stress that when the group G becomes larger and larger the lattice $\text{PM}(G)$ becomes smaller and smaller.

In other words, the theory of permutant measures becomes more and more useful as the groups get bigger and bigger.

Building GENEOS by permutant measures

The previous method for building GENEOS can be generalized by replacing the arithmetic mean with another symmetric function. We can indeed show that when a symmetric function and a *permutant* for the equivariance group G are available, we can easily build a (non-linear) GENEOS with respect to G .

First of all, let us recall the concept of *permutant*, which is related to the one of *permutant measure*.

Definition

We say that a subset $H \subseteq \text{Bij}(X)$ is a *permutant* for G if either $H = \emptyset$ or $gHg^{-1} = H$ for every $g \in G$.

Note that a subset H of $\text{Bij}(X)$ is a permutant for G if and only if H is a union of orbits for the conjugation action of G on $\text{Bij}(X)$.

Building GENEOS by permutant measures

Definition

Let C be a symmetric subset of \mathbb{R}^n , i.e., a subset C such that $\pi(C) = C$ for every permutation π of the coordinates. A function $f: C \rightarrow \mathbb{R}$ is said to be *symmetric on C* if its value is the same no matter the order of its arguments. That is,

$$f(a_1, \dots, a_n) = f(a_{\pi(1)}, \dots, a_{\pi(n)})$$

for every $(a_1, \dots, a_n) \in C$ and every permutation π of the set $\{1, \dots, n\}$.

Building GENEOS by permutant measures

Proposition

Let f be a continuous real-valued symmetric function defined on a compact symmetric subset K of \mathbb{R}^n . Then f is the restriction of a continuous real-valued symmetric function \bar{f} defined on \mathbb{R}^n .

In other words, the concept of continuous real-valued symmetric function defined on a compact symmetric subset K of \mathbb{R}^n coincides with the concept of restriction to K of a continuous real-valued symmetric function defined on \mathbb{R}^n .

Building GENEOS by permutant measures

Let $\mathcal{S} : \mathbb{R}^n \rightarrow \mathbb{R}$ be a symmetric function. If $H = \{h_i\}_{i=1}^n$ is a non-empty permutant for $G \subseteq \text{Bij}_\Phi(X)$, then we can define an operator $\mathcal{S}_H : \Phi \rightarrow \mathbb{R}_b^X$ by setting, for any $\varphi \in \Phi$,

$$\mathcal{S}_H(\varphi) := \mathcal{S}(\varphi \circ h_1, \dots, \varphi \circ h_n),$$

where $\mathcal{S}(\varphi \circ h_1, \dots, \varphi \circ h_n)(x) := \mathcal{S}((\varphi \circ h_1)(x), \dots, (\varphi \circ h_n)(x))$ for every $x \in X$.

Proposition

If $\mathcal{S} : \mathbb{R}^n \rightarrow \mathbb{R}$ is a symmetric function, $G \subseteq \text{Bij}_\Phi(X)$, and H is a non-empty permutant for G , then \mathcal{S}_H is a GEO from (Φ, G) to (\mathbb{R}_b^X, G) with respect to the identity homomorphism $\text{id}_G : G \rightarrow G$. If the restriction of \mathcal{S} to $\text{Im}(\Phi)^n$ is non-expansive, then \mathcal{S}_H is a GENEOS from (Φ, G) to (\mathbb{R}_b^X, G) with respect to id_G .

Building GENEOS by permutant measures

Let us now recall some basic facts about the approximation of symmetric functions by symmetric polynomials. In the sequel, we will denote the symmetric group over the set $\{1, \dots, n\}$ as S_n . Let K be a compact metric space, and $\mathcal{C}(K)$ be the vector space of continuous real-valued functions on K . With a slight abuse of notation, in the following we will confuse each polynomial with the function it represents, restricted to the domain we are considering. Furthermore, if I is a finite subset of \mathbb{N}^n , we will say that a polynomial $\sum_{(k_1, \dots, k_n) \in I} c_{k_1, \dots, k_n} y_1^{k_1} \cdot \dots \cdot y_n^{k_n}$ is a *symmetric polynomial* if $\pi(I) = I$, and $c_{k_1, \dots, k_n} = c_{\pi(k_1), \dots, \pi(k_n)}$ for every multi-index $(k_1, \dots, k_n) \in I$ and every permutation $\pi \in S_n$.

In the next two slides we will show that any symmetric function can be approximated with arbitrary precision by a symmetric polynomial.

Building GENEOS by permutant measures

Definition

A subset A of $\mathcal{C}(K)$ is an *algebra* if it is a vector subspace of $\mathcal{C}(K)$ that is closed under multiplication (i.e., if $f, g \in A$ then $f \cdot g \in A$). Given a set S of functions on K , we say that S *separates points* if for each pair of points $s, t \in K$ there is a function $f \in S$ such that $f(s) \neq f(t)$. Given a set S of functions on K , we say that S *vanishes* at $s \in K$ if $f(s) = 0$ for all $f \in S$.

Theorem (Stone - Weierstrass Theorem)

An algebra A of continuous real-valued functions on a compact metric space K that separates points and does not vanish at any point is dense in $\mathcal{C}(K)$ w.r.t. the max-norm referred to the domain K .

Building GENEOS by permutant measures

Corollary

Let K be a compact subset of \mathbb{R}^n . The algebra of all polynomials $p(y_1, \dots, y_n)$ in n variables is dense in $\mathcal{C}(K)$ with respect to the max-norm referred to the domain K .

Proposition

Let K be a compact and symmetric subset of \mathbb{R}^n . If $\mathcal{S}|_K: K \rightarrow \mathbb{R}$ is the restriction to K of a continuous symmetric function $\mathcal{S}: \mathbb{R}^n \rightarrow \mathbb{R}$ and $\|\cdot\|_\infty$ is the max-norm referred to the domain K , then for every $\varepsilon > 0$ there exists a symmetric polynomial q in n variables such that

$$\|\mathcal{S}|_K - q|_K\|_\infty \leq \varepsilon.$$

Therefore, we can restrict our attention to symmetric polynomials.

Building GENEOS by permutant measures

Definition

The *elementary symmetric polynomials* in the n variables a_1, \dots, a_n , also called *elementary symmetric functions*, are defined as:

$$\sigma_1 := a_1 + \dots + a_n$$

$$\sigma_2 := a_1 \cdot a_2 + a_1 \cdot a_3 + \dots + a_{n-1} \cdot a_n = \sum_{1 \leq i < j \leq n} a_i \cdot a_j$$

⋮

$$\sigma_r := \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} a_{i_1} \cdot a_{i_2} \cdot \dots \cdot a_{i_r} = \sum_{1 \leq i_1 < i_2 < \dots < i_r \leq n} \prod_{j=i_1}^{i_r} a_j$$

⋮

$$\sigma_n := a_1 \cdot a_2 \cdot \dots \cdot a_n.$$

Building GENEOS by permutant measures

We recall an important result in the theory of symmetric polynomials:

Theorem (Fundamental Theorem on Symmetric Polynomials)

Any symmetric polynomial in n variables a_1, \dots, a_n is representable in a unique way as a polynomial in the elementary symmetric polynomials $\sigma_1, \dots, \sigma_n$.

Building GENEOS by permutant measures

In conclusion, if an equivariance group G is chosen and the GEO $F = \mathcal{S}_H$ is built, we can approximate F in the following way, provided that X and Φ are compact. First of all, we can approximate the continuous symmetric function \mathcal{S} by a polynomial $p: \mathbb{R}^n \rightarrow \mathbb{R}$, with an arbitrarily small error ε on the symmetric set $\text{Im}(\Phi)^n$, which is guaranteed to be compact. Then, we can consider the symmetric polynomial $q(a_1, \dots, a_n) := \frac{1}{n!} \sum_{\pi \in S_n} p(a_{\pi(1)}, \dots, a_{\pi(n)})$. Finally, we can consider the GEO F' defined by setting $F'(\varphi) := q(\varphi \circ h_1, \dots, \varphi \circ h_n)$ for every $\varphi \in \Phi$. Since $H \subseteq \text{Bij}_\Phi(X)$, $\|F(\varphi) - F'(\varphi)\|_\infty = \max_{x \in X} |\mathcal{S}(\varphi(h_1(x)), \dots, \varphi(h_n(x))) - q(\varphi(h_1(x)), \dots, \varphi(h_n(x)))| \leq \|S_{|\text{Im}(\Phi)^n} - q_{|\text{Im}(\Phi)^n}\|_\infty = \varepsilon$ for any $\varphi \in \Phi$, and hence the operator F' can be chosen arbitrarily close to F .

Building GENEOS by permutant measures

The Fundamental Theorem on Symmetric Polynomials guarantees that the restriction to $\text{Im}(\Phi)^n$ of any continuous symmetric functions can be approximated arbitrarily well by the restriction to $\text{Im}(\Phi)^n$ of a polynomial in the elementary symmetric functions, defined as

$$\tilde{\mathcal{J}}(a_1, \dots, a_n) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} c_{k_1, \dots, k_n} \prod_{i=1}^n \sigma_i^{k_i}(a_1, \dots, a_n),$$

where $m_i \in \mathbb{N}$ for every $i \in \{1, \dots, n\}$, $c_{k_1, \dots, k_n} \in \mathbb{R}$ for every $k_1 \in \{0, \dots, m_1\}, \dots, k_n \in \{0, \dots, m_n\}$ and σ_i is the i -th elementary symmetric polynomial for every $i \in \{1, \dots, n\}$. We already know that the associated operator is a GEO, but we can indeed obtain a GENEOS by considering a suitable multiple of $\tilde{\mathcal{J}}$.

Building GENEOS by permutant measures

In the sequel, we will need the following constants:

$$M_{\text{Im}(\Phi)^n} := \max_{\alpha \in \text{Im}(\Phi)^n} \|\alpha\|_\infty = \max_{\varphi \in \Phi} \|\varphi\|_\infty \quad (0.1)$$

$$M_1 := \max_{1 \leq i \leq n} \left\{ k_i \binom{n}{i}^{k_i} i M_{\text{Im}(\Phi)^n}^{ik_i-1} \right\} \quad (0.2)$$

$$M_2 := \max_{1 \leq i \leq n} \left\{ \binom{n}{i}^{k_i} M_{\text{Im}(\Phi)^n}^{ik_i} \right\}^{n-1} \quad (0.3)$$

$$C = n \sum_{k_1=0}^{m_1} \cdots \sum_{k_n=0}^{m_n} |c_{k_1, \dots, k_n}| M_1 M_2 \quad (0.4)$$

Building GENEOS by permutant measures

Let us consider a non-empty permutant $H = \{h_i\}_{i=1}^n$ for a subgroup $G \subseteq \text{Bij}_\Phi(X)$. We can define an operator $\hat{\mathcal{J}}_H: \Phi \rightarrow \mathbb{R}_b^X$ by setting

$$\hat{\mathcal{J}}_H(\varphi) := \frac{1}{C} \tilde{\mathcal{J}}(\varphi \circ h_1, \dots, \varphi \circ h_n)$$

for any $\varphi \in \Phi$, where

$\tilde{\mathcal{J}}(\varphi \circ h_1, \dots, \varphi \circ h_n)(x) := \tilde{\mathcal{J}}((\varphi \circ h_1)(x), \dots, (\varphi \circ h_n)(x))$ for every $x \in X$ and C is the constant defined in (0.4).

Building GENEOS by permutant measures

Lemma

For every $\alpha, \beta \in \text{Im}(\Phi)^n$ with $n > 0$, and every $(k_1, \dots, k_n) \in \mathbb{N}^n$

$$\left| \prod_{i=1}^n \sigma_i^{k_i}(\alpha) - \prod_{i=1}^n \sigma_i^{k_i}(\beta) \right| \leq n \|\alpha - \beta\|_\infty M_1 M_2.$$

From the previous lemma the next result follows.

Theorem


Let us assume that $G \subseteq \text{Bij}_\Phi(X)$ and H is a non-empty permutant for G . If $\tilde{\mathcal{J}}$ is a polynomial in the n elementary symmetric functions, then $\hat{\mathcal{J}}_H$ is a GENEOS from (Φ, G) to (\mathbb{R}_b^X, G) with respect to the identity homomorphism $\text{id}_G : G \rightarrow G$.

How can we represent linear GENEOSs?

Annals of Mathematics and Artificial Intelligence

<https://doi.org/10.1007/s10472-022-09830-1>

On the finite representation of linear group equivariant operators via permutant measures

Giovanni Bocchi¹ · Stefano Botteghi² · Martina Brasini² · Patrizio Frosini²  · Nicola Quercioli³

Accepted: 26 December 2022

© The Author(s) 2023

<https://rdcu.be/c5Obw>

Representation Theorem for linear GNEOs

The following theorem strengthens our previous result about building linear GNEOs via permutant measures.

Theorem (Representation Theorem for linear GNEOs)

Let us assume that $G \subseteq \text{Bij}(X)$ transitively acts on the finite set X and that F is a map from \mathbb{R}^X to \mathbb{R}^X . The map F is a linear GNEO from \mathbb{R}^X to \mathbb{R}^X with respect to the identical homomorphism $\text{id}_G: g \mapsto g$ **if and only if** a permutant measure μ with respect to G exists, such that $F(\varphi) = \sum_{h \in \text{Bij}(X)} \varphi h^{-1} \mu(h)$ for every $\varphi \in \mathbb{R}^X$, and $\sum_{h \in \text{Bij}(X)} |\mu(h)| \leq 1$.

Sketch of the proof

Let F be a linear GENE0 from \mathbb{R}^X to \mathbb{R}^X with respect to the identical homomorphism $\text{id}_G: g \mapsto g$.

The proof of the Representation Theorem for linear GENE0s is based on the matrix $B = (b_{ij})$ associated with F with respect to the basis $\{\mathbb{1}_{x_1}, \dots, \mathbb{1}_{x_n}\}$ of the vector space \mathbb{R}^X .

We consider the linear maps $F^\oplus, F^\ominus: \mathbb{R}^X \rightarrow \mathbb{R}^X$ defined by setting $F^\oplus(\mathbb{1}_{x_j}) := \sum_{i=1}^n \max\{b_{ij}, 0\} \mathbb{1}_{x_i}$ and $F^\ominus(\mathbb{1}_{x_j}) := \sum_{i=1}^n \max\{-b_{ij}, 0\} \mathbb{1}_{x_i}$ for every index $j \in \{1, \dots, n\}$. The matrices associated with F^\oplus and F^\ominus with respect to the basis $\{\mathbb{1}_{x_1}, \dots, \mathbb{1}_{x_n}\}$ of \mathbb{R}^X are

$B^\oplus = (b_{ij}^\oplus) = (\max\{b_{ij}, 0\})$ and $B^\ominus = (b_{ij}^\ominus) = (\max\{-b_{ij}, 0\})$, respectively (in particular, B^\oplus, B^\ominus are non-negative matrices);

Sketch of the proof

Since the action of G on X is transitive, we can prove that an n -tuple of real numbers $\alpha = (\alpha_1, \dots, \alpha_n)$ exists such that each row and each column of B^\oplus can be obtained by permuting α . The same holds for B^\ominus . To proceed, we need the following well known Lemma.

Lemma (Birkhoff – von Neumann decomposition)

Let M be a $n \times n$ real matrix with non-negative entries, such that both the sum of the elements of each row and the sum of the elements of each column is equal to \bar{c} . Then for every $h \in \text{Bij}(X)$ a non-negative real number $c(h)$ exists such that $\sum_{h \in \text{Bij}(X)} c(h) = \bar{c}$ and $M = \sum_{h \in \text{Bij}(X)} c(h)P(h)$, where $P(h)$ is the permutation matrix associated with h .

Sketch of the proof

This result follows:

Proposition

For every $h \in \text{Bij}(X)$ two non-negative real numbers $c^\oplus(h), c^\ominus(h)$ exist, such that $F^\oplus(\varphi) = \sum_{h \in \text{Bij}(X)} c^\oplus(h) \varphi h^{-1}$ and $F^\ominus(\varphi) = \sum_{h \in \text{Bij}(X)} c^\ominus(h) \varphi h^{-1}$ for every $\varphi \in \mathbb{R}^X$. Moreover, if $g \in G$ then $B^\oplus P(g) = P(g) B^\oplus$ and $B^\ominus P(g) = P(g) B^\ominus$.

Sketch of the proof

Now, for each $h \in \text{Bij}(X)$, let us consider the orbit $\mathcal{O}(h)$ of h under the conjugation action of G on $\text{Bij}(X)$, and set

$$\begin{aligned}\mu^{\oplus}(h) &:= \sum_{f \in \mathcal{O}(h)} \frac{c^{\oplus}(f)}{|\mathcal{O}(f)|} = \sum_{f \in \mathcal{O}(h)} \frac{c^{\oplus}(f)}{|\mathcal{O}(h)|} \\ \mu^{\ominus}(h) &:= \sum_{f \in \mathcal{O}(h)} \frac{c^{\ominus}(f)}{|\mathcal{O}(f)|} = \sum_{f \in \mathcal{O}(h)} \frac{c^{\ominus}(f)}{|\mathcal{O}(h)|}.\end{aligned}$$

In other words, we define the measures $\mu^{\oplus}(h), \mu^{\ominus}(h)$ of each permutation h as the averages of the functions c^{\oplus}, c^{\ominus} along the orbit of h under the conjugation action of G .

Sketch of the proof

The function $\mu := \mu^{\oplus} - \mu^{\ominus}$ is the permutant measure that we need to prove the statement of the Representation Theorem for linear GNEOs.

The equality $F(\varphi) = \sum_{f \in \text{Bij}(X)} \varphi f^{-1} \mu(f)$ holds, since $F = F^{\oplus} - F^{\ominus}$.

Furthermore, we can prove that

$$\sum_{h \in \text{Bij}(X)} |\mu(h)| = \max_{\varphi \in \mathbb{R}^X \setminus \{0\}} \frac{\|F(\varphi)\|_{\infty}}{\|\varphi\|_{\infty}} \leq 1.$$

The key role of observers in data analysis

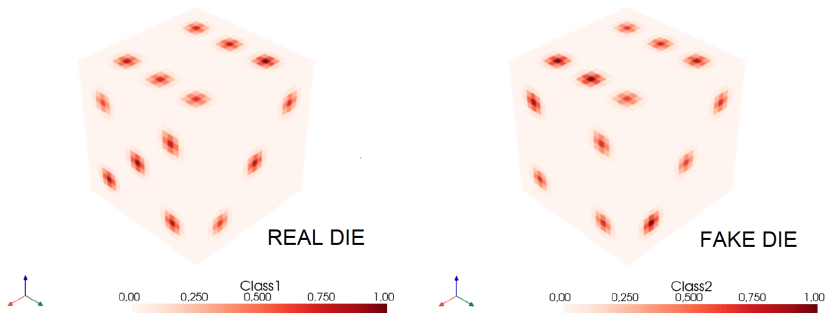
Topological and metric basics for the theory of GENEOS

Building linear and nonlinear GENEOS

How can we use GENEOS in applications?

What happens when we apply GENEOS to our data?

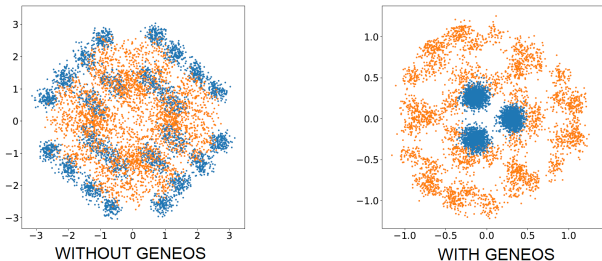
An example of use: comparison between real dice and fake dice.



(Experiment and computations by Giovanni Bocchi)

What happens to data when we apply GENEOS?

We produced 10000 dice (a training set of size 7000 and a test set of size 3000), then we applied PCA to the test set and to the test set transformed by a suitable GENEIO, optimized on the training set:



For each die the first two principal components are plotted. Blue points are associated with **real dice**, while orange ones with **fake dice**. The GENEIO we use was built by a convex combination of 3 GENEIOs defined by permutant measures.

A real application: finding pockets in proteins

GENEOnet: A new machine learning paradigm based on Group Equivariant Non-Expansive Operators. An application to protein pocket detection.

Giovanni Bocchi¹, **Patrizio Frosini**², **Alessandra Micheletti**¹, **Alessandro Pedretti**³
Carmen Gratteri⁴, **Filippo Lunghini**⁵, **Andrea Rosario Beccari**⁵ and **Carmine Talarico**⁵

¹ Department of Environmental Science and Policy, Università degli Studi di Milano

² Department of Mathematics, Università degli Studi di Bologna

³ Department of Pharmaceutical Sciences, Università degli Studi di Milano

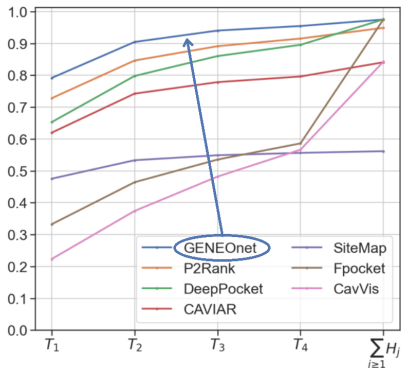
⁴ Dipartimento di Scienze della Salute, Università degli Studi "Magna Græcia di Catanzaro"

⁵ Dompé Farmaceutici SpA

<https://arxiv.org/ftp/arxiv/papers/2202/2202.00451.pdf>

A real application: finding pockets in proteins

Here are the results of our experiments:

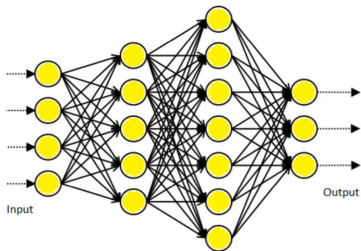


Please note that GENEOnet uses 17 parameters, while a CNN such as DeepPocket requires 665122 parameters.

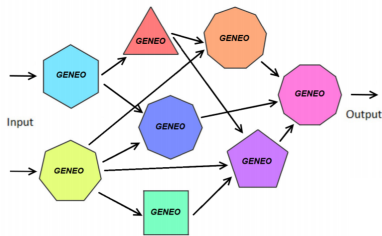
The main point in our approach

In perspective, we are looking for a good compositional theory for building **efficient** and **transparent** networks of GENEOS.

Some preliminary experiments suggest that replacing neurons with GENEOS could make deep learning more transparent and interpretable and speed up the learning process.



NEURAL NETWORK



NETWORK OF GENEOS

THANKS
FOR YOUR
ATTENTION

