Theoretical Underpinnings of Strong Properties

Mathematics
is
the art and science
of
reducing difficult problems
to
linear algebra problems.

–William Stein (see [Z])

In the next two lectures, we:

- discuss how basic results from multi-variable real analysis can be used to implicitly solve various inverse problems;
- illustrate this in the setting of spectrally arbitrary sign-patterns; and
- develop the theoretical underpinnings for one of the strong properties–the Strong Arnol'd Property.

  This will establish a template for other strong properties.

Good news!

Many mathematical questions can be reduced to determining whether or not a system of real-valued functions

$$\begin{aligned} f_1(\mathbf{x}) &= b_1 \\ f_2(\mathbf{x}) &= b_2 \\ &\vdots \\ f_m(\mathbf{x}) &= b_m \end{aligned} \qquad \text{where } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n.$$

has a solution.

Example.

Is there a nilpotent matrix N with sign-pattern

$$S = \begin{bmatrix} - & + & 0 \\ - & 0 & + \\ 0 & - & + \end{bmatrix} ?$$

Equivalent to: are there positive real numbers $a, \ldots, f$ such that

$$N = \begin{bmatrix} -a & e & 0 \\ -b & 0 & f \\ 0 & -c & d \end{bmatrix} \text{ is nilpotent?}$$

Note, by diagonal similarity, may assume that $e = 1$ and $f = 1$.

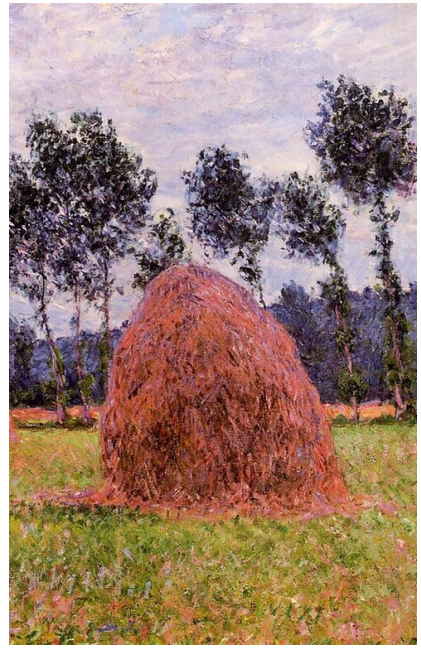
This problem is equivalent to: do there exist positive a, b, c, d such that

$$\begin{aligned} f_1(a,b,c,d) &= a - d &&= 0 \\ f_2(a,b,c,d) &= -ad + b + c &&= 0 \\ f_3(a,b,c,d) &= -ac + bd &&= 0? \end{aligned}$$

Finding a Needle in a Haystack

Bad news

In general, this is a difficult problem–similar to finding a needle in a haystack.



Monet, 1840 (Public domain)

In theory, if each of the f_i is a polynomial, and variables real, then the Real or Postiv Nullstellensatz can be used determine if the system has a solution.

But in superexponential time and space!

Good news!

Often, if there is a needle in a haystack, then we can guarantee that there is a needle in every nearby haystack!



Monet, 1840 (public domain)

That is, often when a system of equations has a solution, then each "nearby" system of equations has a solution!

Example.

$$\begin{aligned} f_1(a,b,c,d) &= a-d &&= 0 \\ f_2(a,b,c,d) &= -ad+b+c &&= 0 \\ f_3(a,b,c,d) &= -ac+bd &&= 0 \\ & a,b,c,d &&> 0? \end{aligned} \tag{1}$$

has solution $(a,b,c,d) = (2,2,2,2)$.

Look at a nearby system of equations

$$\begin{aligned} a-d &= \alpha \\ -ad+b+c &= \beta \\ -ac+bd &= \gamma \end{aligned}$$

where α, β, γ are small.

Example

Look at a nearby system of equations

$$\begin{aligned} a - d &= \alpha \\ -ad + b + c &= \beta \\ -ac + bd &= \gamma \end{aligned}$$

where α, β, γ are small.

This is asking: is there a matrix with sign pattern

$$S = \begin{bmatrix} - & + & 0 \\ - & 0 & + \\ 0 & - & + \end{bmatrix} ?$$

and characteristic polynomial $x^3 + \alpha x^2 + \beta x + \gamma$?

Recall:

> "Mathematics is the art and science of reducing difficult problems to linear algebra problems."

- For a consistent linear system $A\mathbf{x} = \mathbf{b}$ (with A $m \times n$), when is it the case that every nearby system $A\mathbf{x} = \widehat{\mathbf{b}}$ has a solution?

Recall:

> "Mathematics is the art and science of reducing difficult problems to linear algebra problems."

- For a consistent linear system $A\mathbf{x} = \mathbf{b}$ (with A $m \times n$), when is it the case that every nearby system $A\mathbf{x} = \widehat{\mathbf{b}}$ has a solution?

- Exactly when the rank of A is m, or equivalently, when the column space of A is $\mathbb{R}^m$.

Recall:

> "Mathematics is the art and science of reducing difficult problems to linear algebra problems."

- For a consistent linear system $A\mathbf{x} = \mathbf{b}$ (with A $m \times n$), when is it the case that every nearby system $A\mathbf{x} = \widehat{\mathbf{b}}$ has a solution?
- Exactly when the rank of A is m, or equivalently, when the column space of A is $\mathbb{R}^m$.
- How can we reduce the difficult general problem to a linear algebra problem?

Recall:

> "Mathematics is the art and science of reducing difficult problems to linear algebra problems."

- For a consistent linear system $A\mathbf{x} = \mathbf{b}$ (with A $m \times n$), when is it the case that every nearby system $A\mathbf{x} = \widehat{\mathbf{b}}$ has a solution?
- Exactly when the rank of A is m, or equivalently, when the column space of A is $\mathbb{R}^m$.
- How can we reduce the difficult general problem to a linear algebra problem?
- Look at the linearization of f!

Important results and concepts from real variables.

Let $\mathbf{p} \in U$, U be an open subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}^m$ be a function.

f is differentiable at $\mathbf{p}$ provided there exists a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\mathbf{h} \to \mathbf{p}} \frac{f(\mathbf{h}) - f(\mathbf{p}) - L(\mathbf{h} - \mathbf{p})}{\|\mathbf{h} - \mathbf{p}\|} = \mathbf{0}.$$

Here $\| \cdot \|$ is the usual Euclidean norm.

Facts: (see Chapter 2 of Spivak *Calculus on Manifolds*)

- Where f is differentiable at $\mathbf{p}$, L is unique.
- We can view $f = (f_1, \ldots, f_m)$, where $f_i \colon U \to \mathbb{R}$ and $f_i(\mathbf{x}) = f_i(x_1, \ldots, x_n)$.
- If all of the partials $\frac{\partial f_i}{\partial x_j}(\mathbf{p})$ exist and are continuous in a neighborhood of $\mathbf{p}$, then f is differentiable at $\mathbf{p}$.
- Additionally, in this case

$$L(\mathbf{x}) = \left[\frac{\partial f_i}{\partial x_j}(\mathbf{p})\right]\mathbf{x}.$$

- The $m \times n$ matrix $Df(\mathbf{p}) := \left[\frac{\partial f_i}{\partial x_j}(\mathbf{p})\right]$ is the Jacobian of f at $\mathbf{p}$.

Throughout, we abuse notation and use $Df(\mathbf{p})$ to denote both the Jacobian of f and the linearization of f at $\mathbf{p}$.

Example. Let $f : \mathbb{R}^4 \to \mathbb{R}^3$ by

$$\begin{aligned} f_1(a,b,c,d) &= a - d \\ f_2(a,b,c,d) &= -ad + b + c \\ f_3(a,b,c,d) &= -ac + bd \end{aligned}$$

Then

$$Df = \begin{bmatrix} 1 & 0 & 0 & -1 \\ -d & 1 & 1 & -a \\ -c & d & -a & b \end{bmatrix}.$$

Since a and d are positive, $Df(\mathbf{p})$ has rank 3 and the image of $Df(\mathbf{p})$ is all of $\mathbb{R}^3$.

If Stein is correct, f should at least be locally surjective, right?

Local Surjectivity

Theorem (Local Surjectivity Theorem, (see Theorem 2.5.9 of Marsden, et al))

If $f : U \to \mathbb{R}^m$ is continuously differentiable on U and $D_f(\mathbf{u}_0)$ is onto for some $\mathbf{u}_0 \in U$, then f is locally onto.
That is, $\exists$ open neighborhoods U_1 of $\mathbf{u}_0$ and V_1 of $f(\mathbf{u}_0)$ such that $f(U_1) = V_1$.

Notes: $D_f(\mathbf{u}_0)$ is onto if and only if

$$\text{rank} \left[\frac{\partial f_i}{\partial x_j}(\mathbf{u}_0) \right] = m.$$

This is a variant of the Inverse Function theorem.

Back to our example.

$$
\begin{aligned}
f(a,b,c,d) &= (a-d, -ad+b+c, -ac+bd) \\
\mathbf{u}_0 &= (2,2,2,2) \\
f(\mathbf{u}_0) &= \mathbf{0} \\
Df(\mathbf{u}_0) &= \begin{bmatrix} 1 & 0 & 0 & -1 \\ -2 & 1 & 1 & -2 \\ -2 & 2 & -2 & 2 \end{bmatrix}
\end{aligned}
$$

Local Surjectivity theorem $\Rightarrow \exists$ a neighborhood U_1 of $\mathbf{u}_0$ and a neighborhood V_1 of $\mathbf{0}$ such that $f(U_1) = V_1$.

Taking U so that $U \subseteq \mathcal{B}(\mathbf{u}_0, 1)$, we see that for all $\epsilon_1, \epsilon_2, \epsilon_3$ sufficiently small, $\exists$ nonzero a, b, c, d such that $f(a,b,c,d) = (\epsilon_1, \epsilon_2, \epsilon_3)$.

Hence each monic real polynomial $p(x) = x^3 + \epsilon_1 x^2 + \epsilon_2 x + \epsilon_3$ is the characteristic polynomial of a matrix with sign-pattern

$$S = \begin{bmatrix} - & + & 0 \\ - & 0 & + \\ 0 & - & + \end{bmatrix}.$$

Pretty snazzy. We found a special needle in a haystack (the nilpotent matrix) with sign-pattern in S, and used it to show that every nearby haystack has a needle (every monic real polynomial of degree 3 near x^3 is the characteristic polynomial of some matrix with sign-pattern S).

Note by scaling such a matrix by $k > 0$, one can achieve any monic real polynomial of degree 3 as the characteristic polynomial of such a matrix.

Spectrally arbitrary families

A family $\mathcal{F}$ of $n \times n$ matrices is spectrally arbitrary provided each monic, real polynomial of degree n is the characteristic polynomial of at least one matrix in $\mathcal{F}$.

The family of real $n \times n$ companion matrices is spectrally arbitrary.

The family of $n \times n$ diagonal matrices ($n \geq 2$) is not spectrally arbitrary (as one can't realize $(x^2 + 1)x^{n-2}$).

We used the Local Surjectivity Theorem to show the sign-pattern

$$S = \begin{bmatrix} - & + & 0 \\ - & 0 & + \\ 0 & - & + \end{bmatrix}.$$

is spectrally arbitrary.

This is one example of the Nilpotent-Jacobi method developed by Drew, Johnson, Olesky, van den Dreissche (2000).

Kevin will talk much more about this. Here we just introduce the technique.

Nilpotent-Jacobian method

Let $S = [s_{ij}]$ be an $n \times n$ irreducible sign-pattern.

1. Normalize
   $\exists$ a set α of $n-1$ nonzero entries that form a spanning tree of the underlying bipartite graph of S. By sign-similarity we can assume that $s_{ij} = 1$ for the entries in α.

2. Form polynomial map
   Let $S(x)$ be the matrix obtained from S by replacing each nonzero s_{ij} with $(i,j) \notin \alpha$ by the variable $s_{ij}x_{ij}$. Let v be the number of variables.

   We now have a function $f\colon \mathbb{R}^{+v} \to \mathbb{R}^n$ by $\mathbf{x} \mapsto$ the coefficient of λ^i in $\det(\lambda I - A)$ $(i = 0, 1, \ldots, n-1)$.

3. Find a good needle in a haystack
   That is, find a nilpotent matrix with pattern $S(\mathbf{u}_0)$ whose Jacobian matrix $[\partial f_i/\partial x_j(\mathbf{u}_0)]|$ has rank n.

4. Every nearby haystack has a needle.
   That is, every monic real polynomial sufficiently close to x^n is the characteristic polynomial of some matrix with sign-pattern S.

5. S is a spectrally arbitrary sign-pattern!
   To achieve any $p(x)$, achieve with $\epsilon^n p(x/\epsilon)$ with an A and then use $(1/\epsilon)A$.

Big take-aways, this technique is implicit, uses the inverse function theorem, and foreshadows strong properties of matrices.

Kevin's talk will explore this (as well as the inertially arbitrary patterns) in more detail.

Brief aside

As we saw in Lecture 1, there has been significant interest in the maximum nullity of matrices with a given structure; e.g., tridiagonal matrices, Parter and Fiedler for acyclic matrices, Leal-Duarte and Nylen's work on the maximum nullity $M(G)$ of matrices in $\mathcal{S}(G)$.

In the 1990s, Colin de Verdiére developed the strong property and associated parameter for a very special subset of $\mathcal{S}(G)$.

We will talk about this in later lectures. Colin de Verdiére's work was extended to $\mathcal{S}(G)$ with a new parameter by Barioli, Fallat, and Hogben (2005). For simplicity, we develop the BFH theory and parameter first.

Local Surjectivity Theorem & Maximum Multiplicity

Setting:

- n positive integer
- $\mathrm{Sym}(n)$ the set of all $n \times n$ symmetric matrices
- G a graph on n vertices
- $\mathcal{S}(G)$ the set of all symmetric matrices with graph G i.e., $A = [a_{ij}]$ symmetric and for $i \neq j$, $a_{ij} \neq 0$ if and only if ij is an edge of G.
- $\langle \mathcal{S}(G) \rangle$: subspace of $\mathrm{Sym}(G)$ spanned by matrices with graph G.
- $\mathcal{B}(O, \epsilon)$ the set of all matrices A in $\mathrm{Sym}(G)$ with $\|A\| < \epsilon$.

Two natural maps and their derivatives

The perturb nonzero entries map

U an open neighborhood of O in $\langle \mathcal{S}(G) \rangle$.
$f : U \to \mathrm{Sym}(n)$ by $f(B) = A + B$.

Note: If B is sufficiently small, then $f(B) \in \mathcal{S}(G)$.

The local congruence map

$g : \mathcal{B}(O, \epsilon) \to \mathrm{Sym}(n)$ by

$$g(R) = (I + R)^\top A (I + R).$$

Note if ϵ is sufficiently small, $(I + R)$ is invertible, so $g(R)$ is congruent to A; in particular, $g(R)$ and A have the same rank.

Their total derivatives:

$f\colon U \to \mathrm{Sym}(n)$ by $f(B) = A + B$.

$$\frac{f(B) - f(O) - \iota(B)}{\|B\|} = O$$

So $Df(O) = \iota$ is the identity map, and the image of $Df(O)$ is $\langle \mathcal{S}(G) \rangle$.
$g\colon \mathcal{B}(O, \epsilon) \to \mathrm{Sym}(n)$ by

$$g(R) = (I + R)^\top A (I + R).$$

$$\frac{g(R) - g(O)}{\|R\|} = \frac{R^\top A + AR + R^\top AR}{\|R\|}$$

which tends to $\frac{R^\top A + AR}{\|R\|}$ as $R \to O$.
So $Dg(O)$ is the map $R \mapsto R^\top A + AR$.
Image of $Dg(O)$ is $\{R^\top A + AR : R \in \mathbb{R}^{n \times n}\}$.

Combined maps

Two intriguing maps obtained from combining f and g.

$\mathfrak{f} : \mathcal{B}(O, \epsilon) \times \langle \mathcal{S}(G) \rangle \to \operatorname{Sym} n$ by

$$\mathfrak{f}(R, B) = (I + R)^\top A(I + R) + B.$$

$\mathfrak{g} : \mathcal{B}(O, \epsilon) \times \langle \mathcal{S}(G) \rangle \to \operatorname{Sym}(n)$ by

$$\mathfrak{g}(R, B) = (I + R)^\top (A + B)(I + R)$$

Supergraph Theorem for Inertia

What are the implications of $\mathfrak{f}$ being locally surjective?

$\exists$ neighborhood $V \subseteq \mathrm{Sym}(n)$ of A and a neighborhood of U of (O, O) in $\mathcal{B}(O, \epsilon) \times \langle \mathcal{S}(G) \rangle$ such that $\mathfrak{f}(U) = V$.

Take H be a supergraph of G, and E to be the adjacency matrix of $H \setminus G$. For t sufficiently small, $A + tE \in V$. So there exists (R, B) in U with $A + tE = \mathfrak{f}(R, B) = (I + R)^\top A(I + R) + B$.

Note $A + tE - B$ has graph H, and is congruent to A.

Hence, $\exists$ a matrix in $\mathcal{S}(H)$ with the same inertia as A.

Theorem (Supergraph Theorem for Inertia)

$$A \in \mathcal{S}(G) \text{ and } D\mathfrak{f} \text{ is surjective,}$$
$$H \text{ a supergraph of } G$$
$$\Downarrow$$
$$\exists\ B \in \mathcal{S}(H) \text{ with the same inertia as } A.$$

Proximity Theorem for Inertia

What are the implications of $\mathfrak{g}$ being locally surjective?

$\mathfrak{g} : \mathcal{B}(O, \epsilon) \times \langle \mathcal{S}(G) \rangle \to \mathrm{Sym}(n)$ by

$$\mathfrak{g}(R, B) = (I + R)^\top (A + B)(I + R)$$

$\exists$ a neighborhood $V \subseteq \mathrm{Sym}(n)$ of A and a neighborhood of U of (O, O) in $\mathcal{B}(O, \epsilon) \times \langle \mathcal{S}(G) \rangle$ such that $\mathfrak{g}(U) = V$.

Thus, if M is a symmetric matrix sufficiently close to A, then $M = (I + R)^\top (A + B)(I + R)$ for some R and B.

So $((I + R)^{-1})^\top M (I + R)^{-1} = A + B$ is a symmetric matrix in $\mathcal{S}(G)$ with the same inertia as M.

A useful consequence of the proximity property is the following.

Theorem (Bifurcation Theorem for Inertia)

$$A \in \mathcal{S}(G)$$

$$D\mathfrak{g} \textit{ surjective}$$

$$\textit{Inertia of } A \textit{ equal to } (p, q, m)$$

$$m = m_1 + m_2 + m_3, \textit{ and}$$

$$m_1 \geq p \textit{ and } m_2 \geq q,\ m_3 \geq 0$$

$$\Downarrow$$

$$\exists \textit{ a matrix in } \mathcal{S}(G) \textit{ with inertia } (m_1, m_2, m_3).$$

When is $\mathfrak{f}$ (respectively, $\mathfrak{g}$) locally surjective?

By the Locally Surjectivity theorem, sufficient conditions for local surjectivity are:

$$\mathsf{Im}(D\mathfrak{f}(O,O) = \mathrm{Sym}(n)$$

(respectively, when

$$\mathsf{Im}(D\mathfrak{g}(O,O) = \mathrm{Sym}(n)).$$

$\mathfrak{f} : \mathcal{B}(O,\epsilon) \times \langle \mathcal{S}(G) \rangle \to \mathrm{Sym}(n)$ by
$\mathfrak{f}(R,B) = (I+R)^\top A(I+R) + B$
Jacobian of $\mathfrak{f}$ at $(0,0)$ is $\left[\begin{array}{cc} Df(O,O) & Dg(O,O) \end{array} \right]$.

$\mathfrak{g} : \mathcal{B}(O,\epsilon) \times \langle \mathcal{S}(G) \rangle \to \mathrm{Sym}(n)$ by
$\mathfrak{g}(R,B) = (I+R)^\top (A+B)(I+R)$.

Jacobian of $\mathfrak{g}$ at $(0,0)$ is also $\left[\begin{array}{cc} Df(O,O) & Dg(O,O) \end{array} \right]$.

So the images of the linear transformations $D\mathfrak{f}(O,O)$ and $D\mathfrak{g}(O,O)$ are the same and equal

$$\{R^\top A + AR : R \in \mathbb{R}^{n\times n}\} + \langle \mathcal{S}(G)\rangle.$$

This implies that $\mathfrak{f}$ (respectively, $\mathfrak{g}$) is locally surjective when

$$\mathrm{Sym}(n) = \{R^\top A + AR : R \in \mathbb{R}^{n\times n}\} + \langle \mathcal{S}(G)\rangle.$$

Next time, we'll see that this condition is equivalent to the Strong Arnol'd property:

$$\begin{gathered} X \text{ symmetric, } I \circ X = O,\ A \circ X = O \text{ and } AX = O \\ \Downarrow \\ X = O \end{gathered}$$

Local Surjectivity theorem and Intersecting Manifolds

One can view our definition of the SAP from the perspective of manifolds.

Let M_1 and M_2 be manifolds in the same Euclidean space E. Suppose they intersect at a point $\mathbf{p}$.

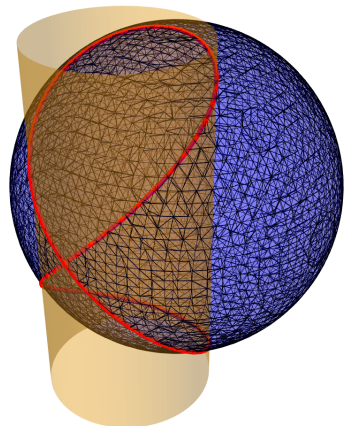
We say that M_1 and M_2 intersect transversally at $\mathbf{p}$ provided the tangent spaces of M_1 and M_2 at $\mathbf{p}$ span the full space E.

Write $M_1 \pitchfork_{\mathbf{p}} M_2$ for M_1 and M_2 intersect transversally at $\mathbf{p}$

Or equivalently, the orthogonal complements of the tangent spaces at $\mathbf{p}$ intersect trivially.

For the SAP, we have two manifolds of interest: the manifold of $n \times n$ matrices of fixed nullity k and the manifold $\rangle \mathcal{S}(G) \langle$

Viviani's curve (see A. Gray)



https://www.desmos.com/3d/vyaoufp2sy

The following well-known theorem could be used for the Super-pattern and the Proximity theorems.

Theorem (Stability of Transversality, (see Guillemin and Pollack))

Let X and W be smooth submanifolds of Z, where X is compact. If X intersects W transversally ($X \pitchfork W$), then there exists a neighborhood $\mathcal{U}$ of the inclusion map $i : X \hookrightarrow Z$ in the C^1 topology such that for every embedding $f \in \mathcal{U}$, the image $f(X)$ also intersects W transversally ($f(X) \pitchfork W$).

References

- F. Barioli, S. M. Fallat and L. Hogben, A variant on the graph parameters of Colin de Verdière: implications to the minimum rank of graphs, Electron. J. Linear Algebra **13** (2005), 387–404.
- S. M. Fallat et al., The bifurcation lemma for strong properties in the inverse eigenvalue problem of a graph, Linear Algebra Appl. **648** (2022), 70–87.
- A. Gray. Viviani's Curve. Section 8.6 in *Modern Differential Geometry of Curves and Surfaces with Mathematica*, 2nd ed. Boca Raton, FL: CRC Press, 1997, pp. 201–202.
- V. Guillemin and A. Pollack. *Differential Topology*. Prentice-Hall, 1974.
- R. H. Abraham, J. E. Marsden and T. S. Ratiu, *Manifolds, tensor analysis, and applications*, second edition, Applied Mathematical Sciences, 75, Springer, New York, 1988.
- M. D. Spivak, *Calculus on manifolds. A modern approach to classical theorems of advanced calculus*, W. A. Benjamin, Inc., New York-Amsterdam, 1965.