The many proofs of the Reduction Phenomenon

Patrick De Leenheer

SPT-CRN 2024

Joint with Lee Altenberg (University of Hawaii) and
Jordan McCaslin (Oregon State University)

June 12, 2024

The Reduction Phenomenon: brief history

- "Increased dispersal (or mutation, or recombination) reduces overall growth."

- First observed in theoretical models by Feldman and colleagues (early '70ies) and called the *Reduction Principle*.

- Proved by Karlin ('76, '82), referred to as *Reduction Phenomenon*.

- Many distinct proofs have been developed over the years, for a wide variety of models (finite and infinite-dimensional, discrete and continuous-time, local and non-local dispersal). In some models, variability in dispersal is captured by a parameter. In other models, two populations with distinct dispersal characteristics (but otherwise equal) are placed in the same environment and we analyze their fate; this is an adaptive dynamics type of approach.

The Reduction Phenomenon: brief history

- This talk focuses on a couple of proofs of the Reduction Phenomenon, for just one of the simplest models.

- What will not be addressed here: If the Reduction Phenomenon holds, it raises an important evolutionary question:

  Why have dispersal strategies evolved?

  Indeed, if increased dispersal lowers overall growth, there should be no advantage to disperse more.

- This suggests that our models, even though they look reasonable, are too simplistic. It calls for including important processes into our model which have thus far been neglected but which may explain why dispersal strategies have evolved.
  [Compare this to the efforts made to modify models exhibiting the Competitive Exclusion Principle or explain the evolution of cooperation etc]

Karlin's Reduction Phenomenon[1]

x_k: n-dimensional vector whose components are the amounts of a species at n spatial locations at time $k = 0, 1, \ldots$.

$$x_{k+1} = [(1-t)I + tP]Dx_k$$

- D: diagonal and $D_{ii} > 0$ for all $i = 1, \ldots, n$. (local growth/decay)
- P: column stochastic ($P^T\mathbf{1} = \mathbf{1}$) and irreducible (dispersal pattern)
- t in $(0, 1)$ (dispersal intensity)

Thm $t \to r(t) := r\left([(1-t)I + tP]D\right)$ is non-increasing.
In fact, $r(t)$ is decreasing, unless $D = aI$, some $a > 0$, and then $r(t) = a$.

[1]For any matrix A, $r(A) := \{|\lambda| \mid \lambda \text{ is e-value of } A\}$ denotes its spectral radius.

Start of proof of Karlin's Thm

Set $A(t) = [(1-t)I + tP]D$, a smooth matrix in t.
By PF and IFT Thm's, we find unique, smooth, positive e-vectors $u(t)$ of $A(t)$ and $v(t)$ of $A^T(t)$, for smooth, positive e-value $r(t)$:

$$\begin{aligned} A(t)u(t) = r(t)u(t) \quad , \quad A^T(t)v(t) &= r(t)v(t), \text{ and} \\ \langle v(t), u(t) \rangle &= 1, \text{ for all } t \text{ in } (0,1). \end{aligned}$$

Differentiating the identity $r(t) = \langle v(t), A(t)u(t) \rangle$ yields:

$$\dot{r}(t) = \langle v(t), \dot{A}(t)u(t) \rangle = \frac{1}{t}(r(t) - \langle v(t), Du(t) \rangle)$$

Done if we can show that for all t in $(0,1)$,

$$r(t) \le \langle v(t), Du(t) \rangle, \text{ and}$$

equality holds in $\le$ only if $D = aI$ for some $a > 0$.

Proof 1: Based on **D**onsker-**V**aradhan-**F**riedland variational formula

Notation:

- $\Sigma = \{x \geq 0 \,|\, \langle \mathbf{1}, x \rangle = 1\}$, all probability vectors.
- For vectors x and $y > 0$, $\frac{x}{y}$ has components x_i/y_i.
- For vectors x and y, $x \circ y$ has components $x_i y_i$.

Thm (DFV variational formula). Let A be non-negative and irreducible, and $u > 0$, $v > 0$ satisfy $Au = r(A)u$, $A^T v = r(A)v$ and $\langle v, u \rangle = 1$. Then

1. $r(A) = \sup_{p \in \Sigma} \inf_{x>0} \langle p, \frac{Ax}{x} \rangle$.
2. Supremum in 1. is achieved for $p = u \circ v$: $r(A) = \inf_{x>0} \langle u \circ v, \frac{Ax}{x} \rangle$.
3. Infimum in 2. is achieved **only for** positive scalar multiples of u.

Proof 1: Back to proof of Karlin's Thm, using DFV formula

Wish to show: $r(t) \le \langle v(t), Du(t)\rangle$, for t in $(0,1)$, and equality holds only if $D = aI$ for some $a > 0$.
By DVF formula, and as $P\pi = \pi$ for some $\pi > 0$,

$$\begin{aligned} r(t) &= \inf_{x>0} \langle u(t) \circ v(t), \frac{[(1-t)I + tP]Dx}{x} \rangle \\ &\le \langle u(t) \circ v(t), \frac{[(1-t)I + tP]D D^{-1}\pi}{D^{-1}\pi} \rangle \\ &= \langle u(t) \circ v(t), \frac{\pi}{D^{-1}\pi} \rangle \\ &= \langle v(t), Du(t) \rangle, \text{ for all } t \text{ in } (0,1) \end{aligned}$$

Equality holds in $\le$ only if $D^{-1}\pi$ is e-vector of $[(1-t)I + tP]D$ for $r(t)$, for all t in $(0,1) \implies \pi = r(t)D^{-1}\pi$, all t in $(0,1)$. Then $r(t) = a$ for some $a > 0$, and $D\pi = a\pi \implies D = aI$.

Proof 2: Based on a linear programming result by Kirkland etal ('06)

Fix $u > 0$ and $v > 0$ with $\langle v, u \rangle = 1$ and set $d = v \circ u$. Define

$$\mathcal{P}(u, v) = \{X \geq 0 \,|\, Xu = u,\ X^T v = v\},\ \text{and}\ f(X) = \langle \mathbf{1}, Xd \rangle.$$

Note: $\mathcal{P}(u, v)$ is $\neq \emptyset$, compact, convex polyhedron, and f is linear.

Lemma (Kirkland, Li & Schreiber '06)

$$f(X) \geq 1,\ \text{for all}\ X\ \text{in}\ \mathcal{P}(u, v),$$

and equality holds in $\geq$ only if X is column stochastic (i.e. $X^T \mathbf{1} = \mathbf{1}$).

Pf (sketch): Let $\mathcal{D} = \{X \geq 0 \,|\, X\mathbf{1} = \mathbf{1} = X^T \mathbf{1}\}$, doubly-stochastic matrices. Define invertible $L : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ by $L(X) = V^{-1} X U^{-1} + I - D^{-1}$. Then one can show:

$$\mathcal{P}(u, v) \subseteq L(\mathcal{D}),\ \text{hence}\ \min_{X \in L(\mathcal{D})} f(X) \leq \min_{X \in \mathcal{P}(u,v)} f(X)$$

Proof 2: Based on linear programming result by Kirkland etal ('06)

Pf (cont.)

$$\min_{X \in L(\mathcal{D})} f(X) \leq \min_{X \in \mathcal{P}(u,v)} f(X)$$

Fact: $\mathcal{D}$ = convex hull of permutation matrices.
(*Birkhoff-von Neumann Thm*)

$\Rightarrow L(\mathcal{D})$ = convex hull of images of permutation matrices under L. For any permutation matrix P, we evaluate

$$f(L(P)) = 1 - n + \sum_{i=1}^{n} \frac{v_{\sigma_i}}{v_i} = 1 + n\left(\frac{1}{n}\sum_{i=1}^{n}\frac{v_{\sigma_i}}{v_i} - 1\right) \geq 1,$$

by the AM-GM inequality. Conclude $1 \leq \min_{X \in \mathcal{P}(u,v)} f(X)$.

To show that equality holds in $\leq$ only if X is column stochastic requires a bit more work; but it only exploits that equality in the AM-GM inequality holds only if all the positive numbers are equal.

Proof 2: Back to proof of Karlin's Thm, using linear programming

Corollary: Let $X \geq 0$, and assume $r(X) > 0$ and $u > 0$ and $v > 0$ satisfy $Xu = r(X)u$, $X^T v = r(X)v$ and $\langle v, u\rangle = 1$. Then

$$r(X) \leq \langle \mathbf{1}, Xv \circ u\rangle,$$

and equality in $\leq$ holds only if all column sums of X equal $r(X)$.

Recall: Wish to show that $r(t) \leq \langle v(t), Du(t)\rangle$, for t in $(0,1)$, and equality holds only if $D = aI$ for some $a > 0$.
For all t in $(0,1)$, the **Corollary** implies that

$$r(t) \leq \langle \mathbf{1}, [(1-t)I + tP]Dv(t) \circ u(t)\rangle = \langle \mathbf{1}, Dv(t) \circ u(t)\rangle = \langle v(t), Du(t)\rangle,$$

and equality holds in $\leq$ only if all column sums of $[(1-t)I + tP]D$ are equal to $r(t)$. The ith column sum is D_{ii}, hence $D = aI$ for some $a > 0$ and $r(t) = a$.

Summary

We presented 2 proofs of Karlin's Reduction Phenomenon.

1. The first one is based on the DVF variational formula, an inherently **nonlinear** result.

2. The second proof is based on a **linear programming** result.
(To be fair, we also use the nonlinear AM-GM inequality, and the Birkhoff-von Neumann Thm.)

Bonus: Proof of the DVF variational formula

Define $\mathcal{B} = \{A \geq 0 \,|\, A\mathbf{1} = A^T\mathbf{1}\}$, all *balanced* non-negative matrices.

Lemma (Friedland & Karlin '75) If $A \in \mathcal{B}$, then

$$\langle \mathbf{1}, \frac{Ax}{x} \rangle \geq \langle \mathbf{1}, A\mathbf{1} \rangle, \text{ for all } x > 0.$$

If in addition, A is irreducible, then equality holds in $\geq$ only if $x = \alpha\mathbf{1}$, for some $\alpha > 0$.

Pf: Since $A \in \mathcal{B}$, there is a diagonal D with $D_{ii} > 0$ for all i, $a > 0$ and $S \in \mathcal{D}$:

$$A + D = aS$$

From this, if Lemma holds for matrices in $\mathcal{D}$, it holds for matrices in $\mathcal{B}$: Indeed, for all $x > 0$,

$$\begin{aligned} \langle \mathbf{1}, \frac{Ax}{x} \rangle &= a\langle \mathbf{1}, \frac{Sx}{x} \rangle - \langle \mathbf{1}, D\mathbf{1} \rangle \\ &\geq a\langle \mathbf{1}, S\mathbf{1} \rangle - \langle \mathbf{1}, D\mathbf{1} \rangle = \langle \mathbf{1}, A\mathbf{1} \rangle \end{aligned}$$

Bonus: Proof of the DVF variational formula

Assume $S \in \mathcal{D}$. The map $x \to \langle \mathbf{1}, \frac{Sx}{x} \rangle$ is homogeneous of degree zero, so we may restrict $x > 0$ to

$$Dom = \left\{ x > 0 \,|\, \prod_{i=1}^{n} x_i = 1 \right\}.$$

For x in Dom, by the AM-GM inequality, and Jensen's inequality,

$$\frac{1}{n}\langle \mathbf{1}, \frac{Sx}{x} \rangle \geq \left(\prod_{i=1}^{n} \frac{(Sx)_i}{x_i} \right)^{1/n} = \left(\prod_{i=1}^{n} (Sx)_i \right)^{1/n} \geq 1, \text{ since}$$

$$\ln \left(\prod_{i=1}^{n} (Sx)_i \right) = \sum_i \ln \left(\sum_j S_{ij} x_j \right) \geq \sum_i \sum_j S_{ij} \ln(x_j) = \sum_j \ln(x_j) = 0.$$

Bonus: Proof of the DVF variational formula

Thus, for $S \in \mathcal{D}$ we proved that $\frac{1}{n}\langle \mathbf{1}, \frac{Sx}{x}\rangle \geq 1$ for all $x > 0$, hence

$$\langle \mathbf{1}, \frac{Sx}{x}\rangle \geq n = \langle \mathbf{1}, S\mathbf{1}\rangle$$

If in addition, $S \in \mathcal{D}$ is irreducible, then equality holds in $\geq$ only if all $\frac{(Sx)_i}{x_i}$ are equal. Then there is some $\beta > 0$ such that $Sx = \beta x$. But as $S \in \mathcal{D}$ is irreducible, the PF Thm implies that $\beta = r(S) = 1$ and that $x = \alpha \mathbf{1}$ for some $\alpha > 0$.

Bonus: Proof of the DVF variational formula

Thm (DFV variational formula). Let A be non-negative and irreducible, and $u > 0$ and $v > 0$ satisfy $Au = r(A)u$, $A^T v = r(A)v$ and $\langle v, u\rangle = 1$. Then

1. $r(A) = \sup_{p\in\Sigma} \inf_{x>0}\langle p, \frac{Ax}{x}\rangle$.
2. Supremum in 1. is achieved for $p = u\circ v$: $r(A) = \inf_{x>0}\langle u\circ v, \frac{Ax}{x}\rangle$.
3. Infimum in 2. is achieved **only for** positive scalar multiples of u.

Pf: For all $p \in \Sigma$, $\inf_{x>0}\langle p, \frac{Ax}{x}\rangle \leq \langle p, \frac{r(A)u}{u}\rangle = \langle p, r(A)\mathbf{1}\rangle = r(A)$

$$\Rightarrow \quad \sup_{p\in\Sigma} \inf_{x>0} \langle p, \frac{Ax}{x}\rangle \leq r(A) \tag{1}$$

Claim: For all $x > 0$,

$$\langle u \circ v, \frac{Ax}{x}\rangle \geq r(A), \tag{2}$$

and equality holds in $\geq$ only if $x = \alpha u$ for some $\alpha > 0$.

Pf of Claim: Note that $VAU\mathbf{1} = r(A)u \circ v = (VAU)^T\mathbf{1}$, so $VAU \in \mathcal{B}$, hence by the **Lemma** of Friedland & Karlin,

$$\langle u \circ v, \frac{Ax}{x} \rangle = \langle \mathbf{1}, \frac{(VAU)x/u}{x/u} \rangle \geq \langle \mathbf{1}, VAU\mathbf{1} \rangle = r(A),$$

and irreducibility of VAU implies that equality holds in $\geq$ only if $x/u = \alpha\mathbf{1}$ for some $\alpha > 0$. That is, only if $x = \alpha u$ for some $\alpha > 0$. We proved items 2 and 3. Let's finish by proving item 1.

Since we get equality in (2) for $x = u$, this implies that

$$\sup_{p \in \Sigma} \inf_{x > 0} \langle p, \frac{Ax}{x} \rangle \geq \inf_{x > 0} \langle u \circ v, \frac{Ax}{x} \rangle = r(A).$$

Together with (1), this inequality establishes the DFV variational formula.

Thank you!