Eigenvalue inequalities, log-convexity and scaling: old results and new applications, a tribute to Sam Karlin

Shmuel Friedland
Univ. Illinois at Chicago

City University of Hong-Kong
10 December, 2010

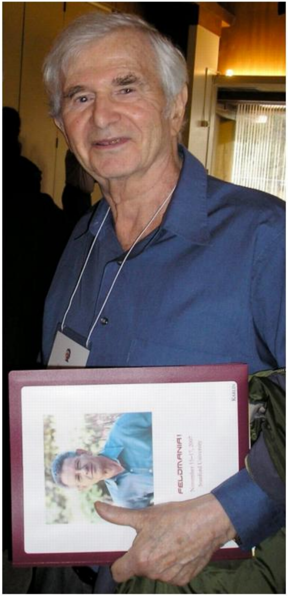

Figure: Karlin

Samuel (Sam) Karlin

Samuel (Sam) Karlin

Karlin was born in Yanovo, Poland, in 1924. He earned his PhD from Princeton as a student of Salomon Bochner in 1947, and was on the faculty at Caltech from 1948 to 1956 before coming to Stanford.

Samuel (Sam) Karlin

Karlin was born in Yanovo, Poland, in 1924. He earned his PhD from Princeton as a student of Salomon Bochner in 1947, and was on the faculty at Caltech from 1948 to 1956 before coming to Stanford.
He made fundamental contributions to game theory, mathematical economics, bioinformatics, probability, evolutionary theory, biomolecular sequence analysis and a field of matrix study known as “total positivity".

Samuel (Sam) Karlin

Karlin was born in Yanovo, Poland, in 1924. He earned his PhD from Princeton as a student of Salomon Bochner in 1947, and was on the faculty at Caltech from 1948 to 1956 before coming to Stanford.
He made fundamental contributions to game theory, mathematical economics, bioinformatics, probability, evolutionary theory, biomolecular sequence analysis and a field of matrix study known as “total positivity".
His main contribution in studying DNA and proteins, was the development (with Amir Dembo and Ofer Zeitouni) of the computer programme BLAST (Basic Local Alignment Search Tool), now the most frequently used software in computational biology.

Samuel (Sam) Karlin

Karlin was born in Yanovo, Poland, in 1924. He earned his PhD from Princeton as a student of Salomon Bochner in 1947, and was on the faculty at Caltech from 1948 to 1956 before coming to Stanford.
He made fundamental contributions to game theory, mathematical economics, bioinformatics, probability, evolutionary theory, biomolecular sequence analysis and a field of matrix study known as “total positivity".
His main contribution in studying DNA and proteins, was the development (with Amir Dembo and Ofer Zeitouni) of the computer programme BLAST (Basic Local Alignment Search Tool), now the most frequently used software in computational biology.
He had 41 doctoral students. He was widely honored: he was a member of the National Academy of Science and the American Academy of Arts and Sciences, and a Foreign Member of the London Mathematical Society. He was the author of 10 books and more than 450 articles.

Samuel (Sam) Karlin

Karlin was born in Yanovo, Poland, in 1924. He earned his PhD from Princeton as a student of Salomon Bochner in 1947, and was on the faculty at Caltech from 1948 to 1956 before coming to Stanford.
He made fundamental contributions to game theory, mathematical economics, bioinformatics, probability, evolutionary theory, biomolecular sequence analysis and a field of matrix study known as “total positivity".
His main contribution in studying DNA and proteins, was the development (with Amir Dembo and Ofer Zeitouni) of the computer programme BLAST (Basic Local Alignment Search Tool), now the most frequently used software in computational biology.
He had 41 doctoral students. He was widely honored: he was a member of the National Academy of Science and the American Academy of Arts and Sciences, and a Foreign Member of the London Mathematical Society. He was the author of 10 books and more than 450 articles.
He died Dec. 18, 2007 at Stanford Hospital after a massive heart

Overview

Overview

- Friedland-Karlin results: Old and New

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975
  - Wielandt 1950 and Donsker-Varadhan 1975

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975
  - Wielandt 1950 and Donsker-Varadhan 1975
  - Log-convexity: Kingman 1961

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975
  - Wielandt 1950 and Donsker-Varadhan 1975
  - Log-convexity: Kingman 1961
  - Diagonal scaling: Sinkhorn 1964, Brualdi-Parter-Schneider 1966 and more

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975
  - Wielandt 1950 and Donsker-Varadhan 1975
  - Log-convexity: Kingman 1961
  - Diagonal scaling: Sinkhorn 1964, Brualdi-Parter-Schneider 1966 and more
- Wireless communication: Friedland-Tan 2008

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975
  - Wielandt 1950 and Donsker-Varadhan 1975
  - Log-convexity: Kingman 1961
  - Diagonal scaling: Sinkhorn 1964, Brualdi-Parter-Schneider 1966 and more
- Wireless communication: Friedland-Tan 2008
  - Statement of the problem

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975
  - Wielandt 1950 and Donsker-Varadhan 1975
  - Log-convexity: Kingman 1961
  - Diagonal scaling: Sinkhorn 1964, Brualdi-Parter-Schneider 1966 and more
- Wireless communication: Friedland-Tan 2008
  - Statement of the problem
  - Relaxation problem

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975
  - Wielandt 1950 and Donsker-Varadhan 1975
  - Log-convexity: Kingman 1961
  - Diagonal scaling: Sinkhorn 1964, Brualdi-Parter-Schneider 1966 and more
- Wireless communication: Friedland-Tan 2008
  - Statement of the problem
  - Relaxation problem
  - SIR domain

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975
  - Wielandt 1950 and Donsker-Varadhan 1975
  - Log-convexity: Kingman 1961
  - Diagonal scaling: Sinkhorn 1964, Brualdi-Parter-Schneider 1966 and more
- Wireless communication: Friedland-Tan 2008
  - Statement of the problem
  - Relaxation problem
  - SIR domain
  - Approximation methods

Overview

- Friedland-Karlin results: Old and New
  - Motivation from population biology
  - Friedland-Karlin 1975
  - Wielandt 1950 and Donsker-Varadhan 1975
  - Log-convexity: Kingman 1961
  - Diagonal scaling: Sinkhorn 1964, Brualdi-Parter-Schneider 1966 and more
- Wireless communication: Friedland-Tan 2008
  - Statement of the problem
  - Relaxation problem
  - SIR domain
  - Approximation methods
  - Direct methods

Motivation from population genetics

Motivation from population genetics

Population is distributed in n demes, subject to local natural selection forces and inter-deme migration pressures

Motivation from population genetics

Population is distributed in n demes, subject to local natural selection forces and inter-deme migration pressures
Migration given by stochastic $A = [a_{ij}] \in \mathbb{R}_+^{n \times n}$

Motivation from population genetics

Population is distributed in n demes, subject to local natural selection forces and inter-deme migration pressures

Migration given by stochastic $A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$

Selection of deme i given by $x_i' = f_i(x_i)$, where

$f_i : [0,1] \to \mathbb{R}_+, \; f_i(0) = f_i(1) = 0, \; f_i'(0) = d_i > 0$

Motivation from population genetics

Population is distributed in n demes, subject to local natural selection forces and inter-deme migration pressures
Migration given by stochastic $A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$
Selection of deme i given by $x_i' = f_i(x_i)$, where
$f_i : [0,1] \to \mathbb{R}_+,\ f_i(0) = f_i(1) = 0,\ f_i'(0) = d_i > 0$

Two possible system of ODE
Either $y_i' = \sum_{j=1}^n a_{ij} f_j(y_j),\ i = 1, \ldots, n$

Motivation from population genetics

Population is distributed in n demes, subject to local natural selection forces and inter-deme migration pressures
Migration given by stochastic $A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$
Selection of deme i given by $x_i' = f_i(x_i)$, where
$f_i : [0,1] \to \mathbb{R}_+,\ f_i(0) = f_i(1) = 0,\ f_i'(0) = d_i > 0$

Two possible system of ODE
Either $y_i' = \sum_{j=1}^n a_{ij} f_j(y_j),\ i = 1,\dots,n$
Or $y_i' = f_i(\sum_{j=1}^n a_{ij} y_j),\ i = 1,\dots,n$

Motivation from population genetics

Population is distributed in n demes, subject to local natural selection forces and inter-deme migration pressures
Migration given by stochastic $A = [a_{ij}] \in \mathbb{R}_+^{n \times n}$
Selection of deme i given by $x_i' = f_i(x_i)$, where
$f_i : [0,1] \to \mathbb{R}_+,\ f_i(0) = f_i(1) = 0,\ f_i'(0) = d_i > 0$

Two possible system of ODE
Either $y_i' = \sum_{j=1}^n a_{ij} f_j(y_j),\ i = 1, \ldots, n$
Or $y_i' = f_i(\sum_{j=1}^n a_{ij} y_j),\ i = 1, \ldots, n$

Find conditions where no species are extinct

Motivation from population genetics

Population is distributed in n demes, subject to local natural selection forces and inter-deme migration pressures
Migration given by stochastic $A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$
Selection of deme i given by $x_i' = f_i(x_i)$, where
$f_i : [0,1] \to \mathbb{R}_+,\ f_i(0) = f_i(1) = 0,\ f_i'(0) = d_i > 0$

Two possible system of ODE
Either $y_i' = \sum_{j=1}^n a_{ij} f_j(y_j),\ i = 1, \ldots, n$
Or $y_i' = f_i(\sum_{j=1}^n a_{ij} y_j),\ i = 1, \ldots, n$

Find conditions where no species are extinct
Linearize at **0** to get iterative system
$\mathbf{z}_j = C\mathbf{z}_{j-1}, j = 1, \ldots$, i.e. $\mathbf{z}_j = C^j \mathbf{z}_0, j = 1, \ldots,$
$C = DA$ Or $C = AD$, $D = \mathrm{diag}(d_1, \ldots, d_n)$

Motivation from population genetics

Population is distributed in n demes, subject to local natural selection forces and inter-deme migration pressures

Migration given by stochastic $A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$

Selection of deme i given by $x_i' = f_i(x_i)$, where

$f_i : [0,1] \to \mathbb{R}_+,\ f_i(0) = f_i(1) = 0,\ f_i'(0) = d_i > 0$

Two possible system of ODE

Either $y_i' = \sum_{j=1}^n a_{ij} f_j(y_j),\ i = 1,\ldots,n$

Or $y_i' = f_i(\sum_{j=1}^n a_{ij} y_j),\ i = 1,\ldots,n$

Find conditions where no species are extinct

Linearize at $\mathbf{0}$ to get iterative system

$\mathbf{z}_j = C\mathbf{z}_{j-1}, j = 1,\ldots,$ i.e. $\mathbf{z}_j = C^j\mathbf{z}_0, j = 1,\ldots,$

$C = DA$ Or $C = AD$, $D = \mathrm{diag}(d_1,\ldots,d_n)$

No species extinct if $\rho(DA) > 1$.

Friedland-Karlin results 1975

Friedland-Karlin results 1975

$A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ irreducible $(I + A)^{n-1} > 0$, $\rho(A)$-spectral radius of A

Friedland-Karlin results 1975

$A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ irreducible $(I + A)^{n-1} > 0$, $\rho(A)$-spectral radius of A
Perron-Frobenius 1907-1908,09,12: STORY (Paul Cohen)

Friedland-Karlin results 1975

$A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ irreducible $(I + A)^{n-1} > 0$, $\rho(A)$-spectral radius of A

Perron-Frobenius 1907-1908,09,12: STORY (Paul Cohen)

$A\mathbf{x}(A) = \rho(A)\mathbf{x}(A),\ \mathbf{x}(A) = (x_1(A), \ldots, x_n(A))^\top > \mathbf{0}$,

$\mathbf{y}(A)^\top A = \rho(A)\mathbf{y}(A),\ \mathbf{y}(A) = (y_1(A), \ldots, y_n(A))^\top > \mathbf{0}$

$\mathbf{x}(A) \circ \mathbf{y}(A) := (x_1(A)y_1(A), \ldots, x_n(A)y_n(A))^\top$-positive probability vector

Friedland-Karlin results 1975

$A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ irreducible $(I + A)^{n-1} > 0$, $\rho(A)$-spectral radius of A
Perron-Frobenius 1907-1908,09,12: STORY (Paul Cohen)
$A\mathbf{x}(A) = \rho(A)\mathbf{x}(A),\ \mathbf{x}(A) = (x_1(A), \ldots, x_n(A))^\top > \mathbf{0}$,
$\mathbf{y}(A)^\top A = \rho(A)\mathbf{y}(A),\ \mathbf{y}(A) = (y_1(A), \ldots, y_n(A))^\top > \mathbf{0}$
$\mathbf{x}(A) \circ \mathbf{y}(A) := (x_1(A)y_1(A), \ldots, x_n(A)y_n(A))^\top$-positive probability vector

THM 1: For $A \geq 0$ irreducible,
$\mathbf{d} = (d_1, \ldots, d_n) > \mathbf{0}$, $D = D(\mathbf{d}) := \mathrm{diag}(d_1, \ldots, d_n)$
$\rho(D(\mathbf{d})A) \geq \rho(A) \prod_{i=1}^n d_i^{x_i(A)y_i(A)}$
If A has positive diagonal then equality holds iff $D(\mathbf{d}) = aI_n$.

Friedland-Karlin results 1975

$A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ irreducible $(I + A)^{n-1} > 0$, $\rho(A)$-spectral radius of A
Perron-Frobenius 1907-1908,09,12: STORY (Paul Cohen)
$A\mathbf{x}(A) = \rho(A)\mathbf{x}(A),\ \mathbf{x}(A) = (x_1(A), \ldots, x_n(A))^\top > \mathbf{0}$,
$\mathbf{y}(A)^\top A = \rho(A)\mathbf{y}(A),\ \mathbf{y}(A) = (y_1(A), \ldots, y_n(A))^\top > \mathbf{0}$
$\mathbf{x}(A) \circ \mathbf{y}(A) := (x_1(A)y_1(A), \ldots, x_n(A)y_n(A))^\top$-positive probability vector

THM 1: For $A \ge 0$ irreducible,
$\mathbf{d} = (d_1, \ldots, d_n) > \mathbf{0}, D = D(\mathbf{d}) := \mathrm{diag}(d_1, \ldots, d_n)$
$\rho(D(\mathbf{d})A) \ge \rho(A)\prod_{i=1}^n d_i^{x_i(A)y_i(A)}$
If A has positive diagonal then equality holds iff $D(\mathbf{d}) = aI_n$.

THM 2: $\min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n x_i(A)y_i(A) \log \frac{(A\mathbf{z})_i}{z_i} = \log \rho(A)$
Equality if $A\mathbf{z} = \rho(A)\mathbf{z}$
If A has positive diagonal then equality iff $A\mathbf{z} = \rho(A)\mathbf{z}$

Friedland-Karlin results 1975

$A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ irreducible $(I+A)^{n-1} > 0$, $\rho(A)$-spectral radius of A
Perron-Frobenius 1907-1908,09,12: STORY (Paul Cohen)
$A\mathbf{x}(A) = \rho(A)\mathbf{x}(A),\ \mathbf{x}(A) = (x_1(A), \ldots, x_n(A))^\top > \mathbf{0},$
$\mathbf{y}(A)^\top A = \rho(A)\mathbf{y}(A),\ \mathbf{y}(A) = (y_1(A), \ldots, y_n(A))^\top > \mathbf{0}$
$\mathbf{x}(A) \circ \mathbf{y}(A) := (x_1(A)y_1(A), \ldots, x_n(A)y_n(A))^\top$-positive probability vector

THM 1: For $A \geq 0$ irreducible,
$\mathbf{d} = (d_1, \ldots, d_n) > \mathbf{0}, D = D(\mathbf{d}) := \mathrm{diag}(d_1, \ldots, d_n)$
$\rho(D(\mathbf{d})A) \geq \rho(A)\prod_{i=1}^n d_i^{x_i(A)y_i(A)}$
If A has positive diagonal then equality holds iff $D(\mathbf{d}) = aI_n$.

THM 2: $\min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n x_i(A)y_i(A)\log\frac{(A\mathbf{z})_i}{z_i} = \log\rho(A)$
Equality if $A\mathbf{z} = \rho(A)\mathbf{z}$
If A has positive diagonal then equality iff $A\mathbf{z} = \rho(A)\mathbf{z}$

COR: $\min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n x_i(A)y_i(A)\frac{(A\mathbf{z})_i}{z_i} = \rho(A)$
(weighted arithmetic-geometric inequality)

Sketch of proofs

THM 2: Assume that A has positive diagonal.

Sketch of proofs

THM 2: Assume that A has positive diagonal.
Restrict $\mathbf{z}$ to probab. vectors Π_n:
$f(\mathbf{z}) = \sum_{i=1}^{n} x_i(A)y_i(A)\log\frac{(A\mathbf{z})_i}{z_i}$

Sketch of proofs

THM 2: Assume that A has positive diagonal.
Restrict $\mathbf{z}$ to probab. vectors Π_n:
$f(\mathbf{z}) = \sum_{i=1}^{n} x_i(A) y_i(A) \log \frac{(A\mathbf{z})_i}{z_i}$
is ∞ on boundary of Π_n

Sketch of proofs

THM 2: Assume that A has positive diagonal.
Restrict $\mathbf{z}$ to probab. vectors Π_n:
$f(\mathbf{z}) = \sum_{i=1}^{n} x_i(A)y_i(A)\log\frac{(A\mathbf{z})_i}{z_i}$
is ∞ on boundary of Π_n
Every critical point in interior Π_n is a strict local minimum

Sketch of proofs

THM 2: Assume that A has positive diagonal.
Restrict $\mathbf{z}$ to probab. vectors Π_n:
$f(\mathbf{z}) = \sum_{i=1}^n x_i(A)y_i(A) \log \frac{(A\mathbf{z})_i}{z_i}$
is ∞ on boundary of Π_n
Every critical point in interior Π_n is a strict local minimum
Hessian of f in $\mathbb{R}^n$ is M-matrix: $H(\mathbf{x}) = \rho(B)I - B$, B symmetric positive

Sketch of proofs

THM 2: Assume that A has positive diagonal.
Restrict $\mathbf{z}$ to probab. vectors Π_n:
$f(\mathbf{z}) = \sum_{i=1}^{n} x_i(A) y_i(A) \log \frac{(A\mathbf{z})_i}{z_i}$
is ∞ on boundary of Π_n
Every critical point in interior Π_n is a strict local minimum
Hessian of f in $\mathbb{R}^n$ is M-matrix: $H(\mathbf{x}) = \rho(B)I - B$, B symmetric positive
Hence unique critical point $\mathbf{x}(A) \in \Pi_n$ is global minimum

Sketch of proofs

THM 2: Assume that A has positive diagonal.
Restrict $\mathbf{z}$ to probab. vectors Π_n:
$f(\mathbf{z}) = \sum_{i=1}^{n} x_i(A)y_i(A)\log\frac{(A\mathbf{z})_i}{z_i}$
is ∞ on boundary of Π_n
Every critical point in interior Π_n is a strict local minimum
Hessian of f in $\mathbb{R}^n$ is M-matrix: $H(\mathbf{x}) = \rho(B)I - B$, B symmetric positive
Hence unique critical point $\mathbf{x}(A) \in \Pi_n$ is global minimum

THM 1: $\rho(DA)\mathbf{x}(DA) = DA\mathbf{x}(DA)$ yields
$\log\rho(DA) = \sum_{i=1}^{n} \mathbf{x}_i(A)\mathbf{y}_i(A)(\log d_i + \frac{(A\mathbf{x}(DA))_i}{x_i(DA)}) \geq$
$\log\rho(A) + \sum_{i=1}^{n} \mathbf{x}_i(A)\mathbf{y}_i(A)\log d_i$

Comparison to Wielandt and Donsker-Varadhan

Comparison to Wielandt and Donsker-Varadhan

$\min_{\mathbf{z}>\mathbf{0}} \max_i \frac{(A\mathbf{z})_i}{z_i} = \max_{\mathbf{z}>\mathbf{0}} \min_i \frac{(A\mathbf{z})_i}{z_i} = \rho(A)$ Wielandt 1950

Comparison to Wielandt and Donsker-Varadhan

$\min_{\mathbf{z}>\mathbf{0}} \max_i \frac{(A\mathbf{z})_i}{z_i} = \max_{\mathbf{z}>\mathbf{0}} \min_i \frac{(A\mathbf{z})_i}{z_i} = \rho(A)$ Wielandt 1950
$\rho(D(\mathbf{d})A) \geq (\max_i d_i)\rho(A)$

Comparison to Wielandt and Donsker-Varadhan

$\min_{\mathbf{z}>\mathbf{0}} \max_i \frac{(A\mathbf{z})_i}{z_i} = \max_{\mathbf{z}>\mathbf{0}} \min_i \frac{(A\mathbf{z})_i}{z_i} = \rho(A)$ Wielandt 1950
$\rho(D(\mathbf{d})A) \ge (\max_i d_i)\rho(A)$

$\max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i = \rho(A)$ Donsker-Varadhan 1975

Comparison to Wielandt and Donsker-Varadhan

$\min_{\mathbf{z}>\mathbf{0}} \max_i \frac{(A\mathbf{z})_i}{z_i} = \max_{\mathbf{z}>\mathbf{0}} \min_i \frac{(A\mathbf{z})_i}{z_i} = \rho(A)$ Wielandt 1950
$\rho(D(\mathbf{d})A) \geq (\max_i d_i)\rho(A)$

$\max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i = \rho(A)$ Donsker-Varadhan 1975

Proof of DV: choose $\boldsymbol{\mu} = \mathbf{x}(A) \circ \mathbf{y}(A)$ use Cor
$\rho(A) \leq \max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i$
$\leq \min_{\mathbf{z}>\mathbf{0}} \max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i = \min_{\mathbf{z}>\mathbf{0}} \max_i \frac{(A\mathbf{z})_i}{z_i} = \rho(A)$

Comparison to Wielandt and Donsker-Varadhan

$\min_{\mathbf{z}>\mathbf{0}} \max_i \frac{(A\mathbf{z})_i}{z_i} = \max_{\mathbf{z}>\mathbf{0}} \min_i \frac{(A\mathbf{z})_i}{z_i} = \rho(A)$ Wielandt 1950
$\rho(D(\mathbf{d})A) \geq (\max_i d_i)\rho(A)$

$\max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i = \rho(A)$ Donsker-Varadhan 1975

Proof of DV: choose $\boldsymbol{\mu} = \mathbf{x}(A) \circ \mathbf{y}(A)$ use Cor
$\rho(A) \leq \max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i$
$\leq \min_{\mathbf{z}>\mathbf{0}} \max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i = \min_{\mathbf{z}>\mathbf{0}} \max_i \frac{(A\mathbf{z})_i}{z_i} = \rho(A)$

Alternatively $\mathbf{z} = e^{\mathbf{u}} = (e^{u_1}, \ldots, e^{u_n})^\top$
$\frac{(Ae^{\mathbf{u}})_i}{e^{u_i}} = \sum_{j=1}^n a_{ij} e^{u_j - u_i}$-convex function

Comparison to Wielandt and Donsker-Varadhan

$\min_{\mathbf{z}>\mathbf{0}} \max_i \frac{(A\mathbf{z})_i}{z_i} = \max_{\mathbf{z}>\mathbf{0}} \min_i \frac{(A\mathbf{z})_i}{z_i} = \rho(A)$ Wielandt 1950
$\rho(D(\mathbf{d})A) \geq (\max_i d_i)\rho(A)$

$\max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i = \rho(A)$ Donsker-Varadhan 1975

Proof of DV: choose $\boldsymbol{\mu} = \mathbf{x}(A) \circ \mathbf{y}(A)$ use Cor
$\rho(A) \leq \max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \min_{\mathbf{z}>\mathbf{0}} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i$
$\leq \min_{\mathbf{z}>\mathbf{0}} \max_{\boldsymbol{\mu}=(\mu_1,\ldots,\mu_n)\in\Pi_n} \sum_{i=1}^n \frac{(A\mathbf{z})_i}{z_i}\mu_i = \min_{\mathbf{z}>\mathbf{0}} \max_i \frac{(A\mathbf{z})_i}{z_i} = \rho(A)$

Alternatively $\mathbf{z} = e^{\mathbf{u}} = (e^{u_1}, \ldots, e^{u_n})^\top$
$\frac{(Ae^{\mathbf{u}})_i}{e^{u_i}} = \sum_{j=1}^n a_{ij} e^{u_j - u_i}$-convex function

$\max_{\boldsymbol{\mu}} \min_{\mathbf{u}\in\mathbb{R}^n} \sum_{i=1}^n \frac{(Ae^{\mathbf{u}})_i}{e^{u_i}}\mu_i = \min_{\mathbf{u}\in\mathbb{R}^n} \max_{\boldsymbol{\mu}} \sum_{i=1}^n \frac{(Ae^{\mathbf{u}})_i}{e^{u_i}}\mu_i$

Convexity of $\rho(A)$ in diagonal entries

Convexity of $\rho(A)$ in diagonal entries

$A_0 = [a_{ij}] \in \mathbb{R}_+^{n \times n}, a_{ii} = 0, i = 1, \ldots, n$, $\mathbf{d} \in \mathbb{R}_+^{n \times n}$

Convexity of $\rho(A)$ in diagonal entries

$A_0 = [a_{ij}] \in \mathbb{R}_+^{n\times n}, a_{ii} = 0, i = 1, \ldots, n$, $\mathbf{d} \in \mathbb{R}_+^{n\times n}$

THM (J.E. Cohen 79): $\rho(A_0 + D(\mathbf{d}))$ is a convex function on $\mathbb{R}_+^n$.

Convexity of $\rho(A)$ in diagonal entries

$A_0 = [a_{ij}] \in \mathbb{R}_+^{n\times n}, a_{ii} = 0, i = 1, \ldots, n$, $\mathbf{d} \in \mathbb{R}_+^{n\times n}$

THM (J.E. Cohen 79): $\rho(A_0 + D(\mathbf{d}))$ is a convex function on $\mathbb{R}_+^n$.

PRF (Friedland 81)
$L(\mathbf{d}, \boldsymbol{\mu}) := \min_{\mathbf{z}>0} \sum_{i=1}^n \frac{((A_0+D(\mathbf{d}))\mathbf{z})_i}{z_i} \mu_i =$
$\sum_{i=1}^n d_i\mu_i + \min_{\mathbf{z}>0} \sum_{i=1}^n \frac{(A_0\mathbf{z})_i}{z_i} \mu_i$
convex in D.

Convexity of $\rho(A)$ in diagonal entries

$A_0 = [a_{ij}] \in \mathbb{R}_+^{n\times n}, a_{ii} = 0, i = 1, \ldots, n$, $\mathbf{d} \in \mathbb{R}_+^{n\times n}$

THM (J.E. Cohen 79): $\rho(A_0 + D(\mathbf{d}))$ is a convex function on $\mathbb{R}_+^n$.

PRF (Friedland 81)
$L(\mathbf{d}, \boldsymbol{\mu}) := \min_{\mathbf{z}>0} \sum_{i=1}^n \frac{((A_0+D(\mathbf{d}))\mathbf{z})_i}{z_i}\mu_i =$
$\sum_{i=1}^n d_i\mu_i + \min_{\mathbf{z}>0} \sum_{i=1}^n \frac{(A_0\mathbf{z})_i}{z_i}\mu_i$
convex in D.

$\rho(D(\mathbf{d}) + A_0) = \max_{\mu\in\Pi_n} L(\mathbf{d}, \mu)$ convex on $\mathbb{R}_+^n$

Convexity of $\rho(A)$ in diagonal entries

$A_0 = [a_{ij}] \in \mathbb{R}_+^{n\times n}, a_{ii} = 0, i = 1, \dots, n$, $\mathbf{d} \in \mathbb{R}_+^{n\times n}$

THM (J.E. Cohen 79): $\rho(A_0 + D(\mathbf{d}))$ is a convex function on $\mathbb{R}_+^n$.

PRF (Friedland 81)
$L(\mathbf{d}, \boldsymbol{\mu}) := \min_{\mathbf{z}>0} \sum_{i=1}^n \frac{((A_0+D(\mathbf{d}))\mathbf{z})_i}{z_i}\mu_i =$
$\sum_{i=1}^n d_i\mu_i + \min_{\mathbf{z}>0} \sum_{i=1}^n \frac{(A_0\mathbf{z})_i}{z_i}\mu_i$
convex in D.

$\rho(D(\mathbf{d}) + A_0) = \max_{\mu\in\Pi_n} L(\mathbf{d}, \mu)$ convex on $\mathbb{R}_+^n$

THM (Friedland 81): $\rho(D(\mathbf{d})A)$ is a convex function on $\mathbb{R}_+^n$ if A^{-1} is M-matrix.

Convexity of $\rho(A)$ in diagonal entries

$A_0 = [a_{ij}] \in \mathbb{R}_+^{n\times n}, a_{ii} = 0, i = 1, \ldots, n,$ $\mathbf{d} \in \mathbb{R}_+^{n\times n}$

THM (J.E. Cohen 79): $\rho(A_0 + D(\mathbf{d}))$ is a convex function on $\mathbb{R}_+^n$.

PRF (Friedland 81)
$L(\mathbf{d}, \boldsymbol{\mu}) := \min_{\mathbf{z}>0} \sum_{i=1}^n \frac{((A_0+D(\mathbf{d}))\mathbf{z})_i}{z_i}\mu_i =$
$\sum_{i=1}^n d_i\mu_i + \min_{\mathbf{z}>0} \sum_{i=1}^n \frac{(A_0\mathbf{z})_i}{z_i}\mu_i$
convex in D.

$\rho(D(\mathbf{d}) + A_0) = \max_{\mu\in\Pi_n} L(\mathbf{d}, \mu)$ convex on $\mathbb{R}_+^n$

THM (Friedland 81): $\rho(D(\mathbf{d})A)$ is a convex function on $\mathbb{R}_+^n$ if A^{-1} is M-matrix.

$A^{-1} = rI - B,\ 0 \le B,\ \rho(B) < r.$

Log-convexity

Log-convexity

A nonnegative function f on convex set $C \subset \mathbb{R}^n$
is log-convex if $\log f$ is convex on C

Log-convexity

A nonnegative function f on convex set $C \subset \mathbb{R}^n$
is log-convex if $\log f$ is convex on C
FACT : The set of log-convex functions on C is a cone closed under multiplication and raising to nonnegative power

Log-convexity

A nonnegative function f on convex set $C \subset \mathbb{R}^n$
is log-convex if $\log f$ is convex on C
FACT : The set of log-convex functions on C is a cone closed under multiplication and raising to nonnegative power

THM (J.F.C. Kingman 1961) $A(\mathbf{x}) = [a_{ij}(\mathbf{x})]_{i=j=1}^{n}$, if each a_{ij} log-convex on C, then $\rho(A(\mathbf{x}))$ log-convex

Log-convexity

A nonnegative function f on convex set $C \subset \mathbb{R}^n$
is log-convex if $\log f$ is convex on C
FACT : The set of log-convex functions on C is a cone closed under multiplication and raising to nonnegative power

THM (J.F.C. Kingman 1961) $A(\mathbf{x}) = [a_{ij}(\mathbf{x})]_{i=j=1}^{n}$, if each a_{ij} log-convex on C, then $\rho(A(\mathbf{x}))$ log-convex

PRF $\rho(A(\mathbf{x})) = \limsup_{m\to\infty}(\mathrm{trace}\ A(\mathbf{x})^m)^{\frac{1}{m}}$

Log-convexity

A nonnegative function f on convex set $C \subset \mathbb{R}^n$
is log-convex if $\log f$ is convex on C
FACT : The set of log-convex functions on C is a cone closed under multiplication and raising to nonnegative power

THM (J.F.C. Kingman 1961) $A(\mathbf{x}) = [a_{ij}(\mathbf{x})]_{i=j=1}^n$, if each a_{ij} log-convex on C, then $\rho(A(\mathbf{x}))$ log-convex

PRF $\rho(A(\mathbf{x})) = \limsup_{m\to\infty}(\mathrm{trace}\ A(\mathbf{x})^m)^{\frac{1}{m}}$

COR: For $A(\mathbf{x})$ as above, $A(\mathbf{x}_0)$ irreducible $\log\rho(A(\mathbf{x})) \geq$
$\log\rho(A(\mathbf{x}_0)) + (\rho(A(\mathbf{x}_0))^{-1}\mathbf{y}(A(\mathbf{x}_0))^\top(\nabla A(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0))\mathbf{x}(A(\mathbf{x}_0))$.

Log-convexity

A nonnegative function f on convex set $C \subset \mathbb{R}^n$
is log-convex if $\log f$ is convex on C
FACT : The set of log-convex functions on C is a cone closed under multiplication and raising to nonnegative power

THM (J.F.C. Kingman 1961) $A(\mathbf{x}) = [a_{ij}(\mathbf{x})]_{i=j=1}^n$, if each a_{ij} log-convex on C, then $\rho(A(\mathbf{x}))$ log-convex

PRF $\rho(A(\mathbf{x})) = \limsup_{m\to\infty}(\mathrm{trace}\ A(\mathbf{x})^m)^{\frac{1}{m}}$

COR: For $A(\mathbf{x})$ as above, $A(\mathbf{x}_0)$ irreducible $\log\rho(A(\mathbf{x})) \ge$
$\log\rho(A(\mathbf{x}_0)) + (\rho(A(\mathbf{x}_0))^{-1}\mathbf{y}(A(\mathbf{x}_0))^\top(\nabla A(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0))\mathbf{x}(A(\mathbf{x}_0))$.

For $A(\mathbf{x}) := D(e^{\mathbf{x}})A$, $A \ge 0$ irreducible $\log\rho(A(\mathbf{x}))$ convex on $\mathbb{R}^n$ and $\mathbf{x}^\top(\mathbf{x}(D(e^{\mathbf{u}})A)\circ\mathbf{y}(D(e^{\mathbf{u}})A))$ is the supporting hyperplane of $\log\rho(A(\mathbf{x}))$ at $\mathbf{u}$

Log-convexity

A nonnegative function f on convex set $C \subset \mathbb{R}^n$
is log-convex if $\log f$ is convex on C
FACT : The set of log-convex functions on C is a cone closed under multiplication and raising to nonnegative power

THM (J.F.C. Kingman 1961) $A(\mathbf{x}) = [a_{ij}(\mathbf{x})]_{i=j=1}^n$, if each a_{ij} log-convex on C, then $\rho(A(\mathbf{x}))$ log-convex

PRF $\rho(A(\mathbf{x})) = \limsup_{m\to\infty} (\mathrm{trace}\, A(\mathbf{x})^m)^{\frac{1}{m}}$

COR: For $A(\mathbf{x})$ as above, $A(\mathbf{x}_0)$ irreducible $\log \rho(A(\mathbf{x})) \ge$
$\log \rho(A(\mathbf{x}_0)) + (\rho(A(\mathbf{x}_0)))^{-1} \mathbf{y}(A(\mathbf{x}_0))^\top (\nabla A(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)) \mathbf{x}(A(\mathbf{x}_0))$.

For $A(\mathbf{x}) := D(e^{\mathbf{x}})A$, $A \ge 0$ irreducible $\log \rho(A(\mathbf{x}))$ convex on $\mathbb{R}^n$ and
$\mathbf{x}^\top(\mathbf{x}(D(e^{\mathbf{u}})A) \circ \mathbf{y}(D(e^{\mathbf{u}})A))$ is the supporting hyperplane of
$\log \rho(A(\mathbf{x}))$ at $\mathbf{u}$
Weaker than Friedland 81 for inverse of M-matrix

Rescaling of irreducible matrices with positive diagonal

Rescaling of irreducible matrices with positive diagonal

THM 3: $A \in \mathbb{R}_+^{n\times n}$ irreducible $\mathbf{0} < \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. If A has positive diagonal then there exists $\mathbf{0} < \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ s.t.

$$D(\mathbf{c})AD(\mathbf{d})\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^\top D(\mathbf{c})AD(\mathbf{d}) = \mathbf{v}^\top$$

$\mathbf{c}, \mathbf{d}$ unique up to $a\mathbf{c}, a^{-1}\mathbf{d}, a > 0$

Rescaling of irreducible matrices with positive diagonal

THM 3: $A \in \mathbb{R}_+^{n\times n}$ irreducible $\mathbf{0} < \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. If A has positive diagonal then there exists $\mathbf{0} < \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ s.t.

$$D(\mathbf{c})AD(\mathbf{d})\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^\top D(\mathbf{c})AD(\mathbf{d}) = \mathbf{v}^\top$$

$\mathbf{c}, \mathbf{d}$ unique up to $a\mathbf{c}, a^{-1}\mathbf{d}, a > 0$

PROOF: $\mathbf{w} = (w_1, \ldots, w_n) := \mathbf{u} \circ \mathbf{v}$. Then $f_{\mathbf{w}}(\mathbf{z}) := \sum_{i=1}^n w_i \log \frac{(A\mathbf{z})_i}{z_i}$ on Π_n has unique critical point in interior of Π_n

Rescaling of irreducible matrices with positive diagonal

THM 3: $A \in \mathbb{R}_+^{n\times n}$ irreducible $\mathbf{0} < \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. If A has positive diagonal then there exists $\mathbf{0} < \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ s.t.

$$D(\mathbf{c})AD(\mathbf{d})\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^\top D(\mathbf{c})AD(\mathbf{d}) = \mathbf{v}^\top$$

$\mathbf{c}, \mathbf{d}$ unique up to $a\mathbf{c}, a^{-1}\mathbf{d}, a > 0$

PROOF: $\mathbf{w} = (w_1, \ldots, w_n) := \mathbf{u} \circ \mathbf{v}$. Then $f_{\mathbf{w}}(\mathbf{z}) := \sum_{i=1}^n w_i \log \frac{(A\mathbf{z})_i}{z_i}$ on Π_n has unique critical point in interior of Π_n
Equivalently $f_{\mathbf{w}}(e^{\mathbf{u}})$ is strictly convex on $\{\mathbf{u} \in \mathbb{R}^n : \mathbf{1}^\top \mathbf{u} = 0\}$.

Rescaling of irreducible matrices with positive diagonal

THM 3: $A \in \mathbb{R}_+^{n\times n}$ irreducible $\mathbf{0} < \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. If A has positive diagonal then there exists $\mathbf{0} < \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ s.t.

$$D(\mathbf{c})AD(\mathbf{d})\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^\top D(\mathbf{c})AD(\mathbf{d}) = \mathbf{v}^\top$$

$\mathbf{c}, \mathbf{d}$ unique up to $a\mathbf{c}, a^{-1}\mathbf{d}, a > 0$

PROOF: $\mathbf{w} = (w_1, \ldots, w_n) := \mathbf{u} \circ \mathbf{v}$. Then $f_{\mathbf{w}}(\mathbf{z}) := \sum_{i=1}^n w_i \log \frac{(A\mathbf{z})_i}{z_i}$ on Π_n has unique critical point in interior of Π_n
Equivalently $f_{\mathbf{w}}(e^{\mathbf{u}})$ is strictly convex on $\{\mathbf{u} \in \mathbb{R}^n : \mathbf{1}^\top \mathbf{u} = 0\}$.

Example 1: $A = \begin{bmatrix} * & * \\ * & 0 \end{bmatrix}$ is not a pattern of doubly stochastic matrix

Rescaling of irreducible matrices with positive diagonal

THM 3: $A \in \mathbb{R}_+^{n\times n}$ irreducible $\mathbf{0} < \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. If A has positive diagonal then there exists $\mathbf{0} < \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ s.t.

$$D(\mathbf{c})AD(\mathbf{d})\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^\top D(\mathbf{c})AD(\mathbf{d}) = \mathbf{v}^\top$$

$\mathbf{c}, \mathbf{d}$ unique up to $a\mathbf{c}, a^{-1}\mathbf{d}, a > 0$

PROOF: $\mathbf{w} = (w_1, \ldots, w_n) := \mathbf{u} \circ \mathbf{v}$. Then $f_{\mathbf{w}}(\mathbf{z}) := \sum_{i=1}^n w_i \log \frac{(A\mathbf{z})_i}{z_i}$ on Π_n has unique critical point in interior of Π_n
Equivalently $f_{\mathbf{w}}(e^{\mathbf{u}})$ is strictly convex on $\{\mathbf{u} \in \mathbb{R}^n : \mathbf{1}^\top \mathbf{u} = 0\}$.

Example 1: $A = \begin{bmatrix} * & * \\ * & 0 \end{bmatrix}$ is not a pattern of doubly stochastic matrix

Example 2: $A = \begin{bmatrix} 0 & * \\ * & 0 \end{bmatrix}$ always rescalable to doubly stochastic with many more solutions than in THM 3.

Fully indecomposable matrices

Fully indecomposable matrices

$A \in \mathbb{R}_+^{n\times n}$ is fully indecomposable, FI, if A has no $s \times (n-s)$ zero submatrix for some integer $s \in [1, n-1]$

Fully indecomposable matrices

$A \in \mathbb{R}_+^{n\times n}$ is fully indecomposable, FI, if A has no $s \times (n-s)$ zero submatrix for some integer $s \in [1, n-1]$

Brualdi-Parter-Schneider 1966: A FI iff PAQ irreducible and has positive diagonal for some permutation matrices P, Q.

Fully indecomposable matrices

$A \in \mathbb{R}_+^{n\times n}$ is fully indecomposable, FI, if A has no $s \times (n-s)$ zero submatrix for some integer $s \in [1, n-1]$

Brualdi-Parter-Schneider 1966: A FI iff PAQ irreducible and has positive diagonal for some permutation matrices P, Q.

Use Frobenius-König THM

Fully indecomposable matrices

$A \in \mathbb{R}_+^{n\times n}$ is fully indecomposable, FI, if A has no $s \times (n-s)$ zero submatrix for some integer $s \in [1, n-1]$
Brualdi-Parter-Schneider 1966: A FI iff PAQ irreducible and has positive diagonal for some permutation matrices P, Q.
Use Frobenius-König THM

THM (BPS66, Sinkhorn-Knopp 67) FI A is diagonally equivalent to doubly stochastic, i.e $D_1 A D_2$ d.s., D_1, D_2 unique up to scalar rescaling

Fully indecomposable matrices

$A \in \mathbb{R}_+^{n\times n}$ is fully indecomposable, FI, if A has no $s \times (n-s)$ zero submatrix for some integer $s \in [1, n-1]$
Brualdi-Parter-Schneider 1966: A FI iff PAQ irreducible and has positive diagonal for some permutation matrices P, Q.
Use Frobenius-König THM

THM (BPS66, Sinkhorn-Knopp 67) FI A is diagonally equivalent to doubly stochastic, i.e $D_1 A D_2$ d.s., D_1, D_2 unique up to scalar rescaling

PRF $D_3(PAQ)D_4\mathbf{1} = (D_3(PAQ)D_4)^\top \mathbf{1} = \mathbf{1}$,
i.e. $D_3 PAQD_4$ d.s.
hence $(P^\top D_3 P)A(QD_4 Q^\top)$ d.s.

Fully indecomposable matrices

$A \in \mathbb{R}_+^{n\times n}$ is fully indecomposable, FI, if A has no $s \times (n-s)$ zero submatrix for some integer $s \in [1, n-1]$
Brualdi-Parter-Schneider 1966: A FI iff PAQ irreducible and has positive diagonal for some permutation matrices P, Q.
Use Frobenius-König THM

THM (BPS66, Sinkhorn-Knopp 67) FI A is diagonally equivalent to doubly stochastic, i.e D_1AD_2 d.s., D_1, D_2 unique up to scalar rescaling

PRF $D_3(PAQ)D_4\mathbf{1} = (D_3(PAQ)D_4)^\top\mathbf{1} = \mathbf{1}$,
i.e. D_3PAQD_4 d.s.
hence $(P^\top D_3P)A(QD_4Q^\top)$ d.s.

Contrary to claim in FK75 I do not know how to prove THM3 for FI matrices

Fully indecomposable matrices

$A \in \mathbb{R}_+^{n\times n}$ is fully indecomposable, FI, if A has no $s \times (n-s)$ zero submatrix for some integer $s \in [1, n-1]$
Brualdi-Parter-Schneider 1966: A FI iff PAQ irreducible and has positive diagonal for some permutation matrices P, Q.
Use Frobenius-König THM

THM (BPS66, Sinkhorn-Knopp 67) FI A is diagonally equivalent to doubly stochastic, i.e D_1AD_2 d.s., D_1, D_2 unique up to scalar rescaling

PRF $D_3(PAQ)D_4\mathbf{1} = (D_3(PAQ)D_4)^\top\mathbf{1} = \mathbf{1}$,
i.e. D_3PAQD_4 d.s.
hence $(P^\top D_3P)A(QD_4Q^\top)$ d.s.

Contrary to claim in FK75 I do not know how to prove THM3 for FI matrices
Reason: why $f(\mathbf{z}) := \sum_{i=1}^n w_i \log \frac{(A\mathbf{z})_i}{z_i}$ blows to ∞ on $\partial\Pi_n$, or attains minimum in the interior of Π_n?

Irreducible matrices with zero diagonal entries - FT08

Irreducible matrices with zero diagonal entries - FT08

THM: $\exists\ A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ has positive off-diagonal entries.
$\mathbf{0} < \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ given. $\mathbf{w} = (w_1, \ldots, w_n) = \mathbf{u} \circ v$. There exists
$\mathbf{0} < \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ s.t.

$$D(\mathbf{c})AD(\mathbf{d})\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^\top D(\mathbf{c})AD(\mathbf{d}) = \mathbf{v}^\top$$

(SC): if $w_i < \sum_{j\neq i} w_j$ for each $a_{ii} = 0$
$\mathbf{c}, \mathbf{d}$ unique up scalar scaling

Irreducible matrices with zero diagonal entries - FT08

THM: 3 $A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ has positive off-diagonal entries.
$\mathbf{0} < \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ given. $\mathbf{w} = (w_1, \ldots, w_n) = \mathbf{u} \circ v$. There exists
$\mathbf{0} < \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ s.t.

$$D(\mathbf{c})AD(\mathbf{d})\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^\top D(\mathbf{c})AD(\mathbf{d}) = \mathbf{v}^\top$$

(SC): if $w_i < \sum_{j\neq i} w_j$ for each $a_{ii} = 0$
$\mathbf{c}, \mathbf{d}$ unique up scalar scaling

THM: 4 $A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ has positive off-diagonal entries, $\mathbf{0} < \mathbf{w} \in \Pi_n$.
Assume (SC) Then
$\max_{\mathbf{z}>0} \sum_{i=1}^n w_i \log \frac{z_i}{(A\mathbf{z})_i} = \sum_{i=1}^n w_i \log(c_i d_i)$,
where $\mathbf{u} = (1, \ldots, 1)^\top, \mathbf{v} = \mathbf{w}$ and $\mathbf{c}, \mathbf{d}$ are given in THM 3.

Irreducible matrices with zero diagonal entries - FT08

THM: 3 $A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ has positive off-diagonal entries. $\mathbf{0} < \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ given. $\mathbf{w} = (w_1, \ldots, w_n) = \mathbf{u} \circ v$. There exists $\mathbf{0} < \mathbf{c}, \mathbf{d} \in \mathbb{R}^n$ s.t.

$$D(\mathbf{c})AD(\mathbf{d})\mathbf{u} = \mathbf{u}, \quad \mathbf{v}^\top D(\mathbf{c})AD(\mathbf{d}) = \mathbf{v}^\top$$

(SC): if $w_i < \sum_{j\neq i} w_j$ for each $a_{ii} = 0$
$\mathbf{c}, \mathbf{d}$ unique up scalar scaling

THM: 4 $A = [a_{ij}] \in \mathbb{R}_+^{n\times n}$ has positive off-diagonal entries, $\mathbf{0} < \mathbf{w} \in \Pi_n$.
Assume (SC) Then
$\max_{\mathbf{z}>0} \sum_{i=1}^n w_i \log \frac{z_i}{(A\mathbf{z})_i} = \sum_{i=1}^n w_i \log(c_i d_i)$,
where $\mathbf{u} = (1, \ldots, 1)^\top, \mathbf{v} = \mathbf{w}$ and $\mathbf{c}, \mathbf{d}$ are given in THM 3.

Proof:
$\sum_{i=1}^n w_i \log \frac{d_i y_i}{(AD(\mathbf{d})\mathbf{y})_i} = \sum_{i=1}^n w_i \log \frac{y_i}{(D(\mathbf{c})AD(\mathbf{d})\mathbf{y})_i} + \sum_{i=1}^n w_i \log(c_i d_i)$

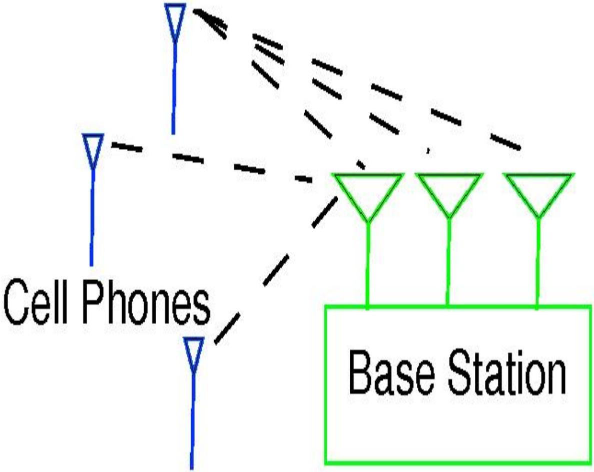

Figure: Cell phones communication

Wireless communication- Statement of the problem

Wireless communication- Statement of the problem

n wireless users. Each transmits with power $p_i \in [0, \bar{p}_i]$, which can be regulated

Wireless communication- Statement of the problem

n wireless users. Each transmits with power $p_i \in [0, \bar{p}_i]$, which can be regulated

$$\mathbf{p} = (p_1, \ldots, p_n) \geq \mathbf{0},\ \bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_n)^\top > \mathbf{0}, \boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^\top > \mathbf{0}$$

Wireless communication- Statement of the problem

n wireless users. Each transmits with power $p_i \in [0, \bar{p}_i]$, which can be regulated

$\mathbf{p} = (p_1, \ldots, p_n) \geq \mathbf{0}, \ \bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_n)^\top > \mathbf{0}, \boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^\top > \mathbf{0}$

Signal-to-Interference Ratio (SIR): $\gamma_i(\mathbf{p}) := \frac{g_{ii}p_i}{\sum_{j\neq i} g_{ij}p_j + \nu_j}$

g_{ii}-amplification, ν_i-AWGN power, $g_{ij}p_j$-interference due to transmitter j

Wireless communication- Statement of the problem

n wireless users. Each transmits with power $p_i \in [0, \bar{p}_i]$, which can be regulated

$\mathbf{p} = (p_1, \ldots, p_n) \geq \mathbf{0}, \ \bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_n)^\top > \mathbf{0}, \boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^\top > \mathbf{0}$

Signal-to-Interference Ratio (SIR): $\gamma_i(\mathbf{p}) := \frac{g_{ii}p_i}{\sum_{j \neq i} g_{ij}p_j + \nu_j}$

g_{ii}-amplification, ν_i-AWGN power, $g_{ij}p_j$-interference due to transmitter j

$\gamma(\mathbf{p}) = (\gamma_1(\mathbf{p}), \ldots, \gamma_n(\mathbf{p}))^\top$

Wireless communication- Statement of the problem

n wireless users. Each transmits with power $p_i \in [0, \bar{p}_i]$, which can be regulated

$\mathbf{p} = (p_1, \ldots, p_n) \geq \mathbf{0},\ \bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_n)^\top > \mathbf{0}, \boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^\top > \mathbf{0}$
Signal-to-Interference Ratio (SIR): $\gamma_i(\mathbf{p}) := \frac{g_{ii}p_i}{\sum_{j\neq i} g_{ij}p_j + \nu_j}$
g_{ii}-amplification, ν_i-AWGN power, $g_{ij}p_j$-interference due to transmitter j

$\gamma(\mathbf{p}) = (\gamma_1(\mathbf{p}), \ldots, \gamma_n(\mathbf{p}))^\top$
$\Phi_{\mathbf{w}}(\gamma) := \sum_{i=1}^{n} w_i \log(1 + \gamma_i),\ \gamma \geq \mathbf{0},\ \mathbf{w} \in \Pi_n$

Wireless communication- Statement of the problem

n wireless users. Each transmits with power $p_i \in [0, \bar{p}_i]$, which can be regulated

$\mathbf{p} = (p_1, \ldots, p_n) \geq \mathbf{0},\ \bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_n)^\top > \mathbf{0}, \boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^\top > \mathbf{0}$
Signal-to-Interference Ratio (SIR): $\gamma_i(\mathbf{p}) := \frac{g_{ii} p_i}{\sum_{j \neq i} g_{ij} p_j + \nu_j}$
g_{ii}-amplification, ν_i-AWGN power, $g_{ij} p_j$-interference due to transmitter j

$\gamma(\mathbf{p}) = (\gamma_1(\mathbf{p}), \ldots, \gamma_n(\mathbf{p}))^\top$
$\Phi_{\mathbf{w}}(\gamma) := \sum_{i=1}^n w_i \log(1 + \gamma_i),\ \gamma \geq \mathbf{0},\ \mathbf{w} \in \Pi_n$

Maximizing sum rates in Gaussian interference-limited channel

$$\max_{\mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}}} \sum_{i=1}^n w_i \log(1 + \gamma_i(\mathbf{p})) = \max_{\mathbf{0} \leq \mathbf{p} \leq \bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\gamma(\mathbf{p})) = \Phi_{\mathbf{w}}(p^\star)$$

Wireless communication- Statement of the problem

n wireless users. Each transmits with power $p_i \in [0, \bar{p}_i]$, which can be regulated

$\mathbf{p} = (p_1, \ldots, p_n) \geq \mathbf{0}$, $\bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_n)^\top > \mathbf{0}, \boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^\top > \mathbf{0}$
Signal-to-Interference Ratio (SIR): $\gamma_i(\mathbf{p}) := \frac{g_{ii}p_i}{\sum_{j\neq i} g_{ij}p_j + \nu_j}$
g_{ii}-amplification, ν_i-AWGN power, $g_{ij}p_j$-interference due to transmitter j

$\gamma(\mathbf{p}) = (\gamma_1(\mathbf{p}), \ldots, \gamma_n(\mathbf{p}))^\top$
$\Phi_{\mathbf{w}}(\gamma) := \sum_{i=1}^n w_i \log(1 + \gamma_i)$, $\gamma \geq \mathbf{0}$, $\mathbf{w} \in \Pi_n$

Maximizing sum rates in Gaussian interference-limited channel

$$\max_{\mathbf{0}\leq\mathbf{p}\leq\bar{\mathbf{p}}} \sum_{i=1}^{n} w_i \log(1 + \gamma_i(\mathbf{p})) = \max_{\mathbf{0}\leq\mathbf{p}\leq\bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\gamma(\mathbf{p})) = \Phi_{\mathbf{w}}(p^\star)$$

Equivalent to maximazing convex function on unbounded convex domain

Wireless communication- Statement of the problem

n wireless users. Each transmits with power $p_i \in [0, \bar{p}_i]$, which can be regulated

$\mathbf{p} = (p_1, \ldots, p_n) \geq \mathbf{0},\ \bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_n)^\top > \mathbf{0}, \boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^\top > \mathbf{0}$
Signal-to-Interference Ratio (SIR): $\gamma_i(\mathbf{p}) := \frac{g_{ii}p_i}{\sum_{j\neq i} g_{ij}p_j + \nu_j}$
g_{ii}-amplification, ν_i-AWGN power, $g_{ij}p_j$-interference due to transmitter j

$\gamma(\mathbf{p}) = (\gamma_1(\mathbf{p}), \ldots, \gamma_n(\mathbf{p}))^\top$
$\Phi_{\mathbf{w}}(\gamma) := \sum_{i=1}^n w_i \log(1 + \gamma_i),\ \gamma \geq \mathbf{0},\ \mathbf{w} \in \Pi_n$

Maximizing sum rates in Gaussian interference-limited channel

$$\max_{\mathbf{0}\leq\mathbf{p}\leq\bar{\mathbf{p}}} \sum_{i=1}^n w_i \log(1 + \gamma_i(\mathbf{p})) = \max_{\mathbf{0}\leq\mathbf{p}\leq\bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\gamma(\mathbf{p})) = \Phi_{\mathbf{w}}(p^\star)$$

Equivalent to maximazing convex function on unbounded convex domain Use for Approximation and Direct methods

Relaxation problem

Relaxation problem

For $\mathbf{z} = (z_1, \ldots, z_n)^\top > \mathbf{0}$ let $\mathbf{z}^{-1} := (z_1^{-1}, \ldots, z_n^{-1})^\top$

Relaxation problem

For $\mathbf{z} = (z_1, \ldots, z_n)^\top > \mathbf{0}$ let $\mathbf{z}^{-1} := (z_1^{-1}, \ldots, z_n^{-1})^\top$
$\gamma(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p} + \boldsymbol{\mu})^{-1}$, $\boldsymbol{\mu} = (\frac{\nu_1}{g_{11}}, \ldots, \frac{\nu_n}{g_{nn}})^\top$

Relaxation problem

For $\mathbf{z} = (z_1, \ldots, z_n)^\top > \mathbf{0}$ let $\mathbf{z}^{-1} := (z_1^{-1}, \ldots, z_n^{-1})^\top$
$\gamma(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p} + \boldsymbol{\mu})^{-1}$, $\boldsymbol{\mu} = (\frac{\nu_1}{g_{11}}, \ldots, \frac{\nu_n}{g_{nn}})^\top$
$F = [f_{ij}] \in \mathbb{R}_+^{n\times n}$ has zero diagonal and $f_{ij} = \frac{g_{ij}}{g_{ii}}$ for $i \neq j$

Relaxation problem

For $\mathbf{z} = (z_1, \ldots, z_n)^\top > \mathbf{0}$ let $\mathbf{z}^{-1} := (z_1^{-1}, \ldots, z_n^{-1})^\top$
$\gamma(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p} + \boldsymbol{\mu})^{-1}$, $\boldsymbol{\mu} = (\frac{\nu_1}{g_{11}}, \ldots, \frac{\nu_n}{g_{nn}})^\top$
$F = [f_{ij}] \in \mathbb{R}_+^{n \times n}$ has zero diagonal and $f_{ij} = \frac{g_{ij}}{g_{ii}}$ for $i \neq j$

$\gamma_{nls}(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p})^{-1}$

Relaxation problem

For $\mathbf{z} = (z_1, \ldots, z_n)^\top > \mathbf{0}$ let $\mathbf{z}^{-1} := (z_1^{-1}, \ldots, z_n^{-1})^\top$
$\gamma(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p} + \boldsymbol{\mu})^{-1}$, $\boldsymbol{\mu} = (\frac{\nu_1}{g_{11}}, \ldots, \frac{\nu_n}{g_{nn}})^\top$
$F = [f_{ij}] \in \mathbb{R}_+^{n\times n}$ has zero diagonal and $f_{ij} = \frac{g_{ij}}{g_{ii}}$ for $i \neq j$

$\gamma_{nls}(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p})^{-1}$

$\Phi_{\mathbf{w},rel}(\gamma) := \sum_{i=1}^n w_i \log \gamma_i, \ \gamma > \mathbf{0}$
obtained by replacing $\log(1+t)$ with smaller $\log t$

Relaxation problem

For $\mathbf{z} = (z_1, \ldots, z_n)^\top > \mathbf{0}$ let $\mathbf{z}^{-1} := (z_1^{-1}, \ldots, z_n^{-1})^\top$
$\gamma(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p} + \boldsymbol{\mu})^{-1}$, $\boldsymbol{\mu} = (\frac{\nu_1}{g_{11}}, \ldots, \frac{\nu_n}{g_{nn}})^\top$
$F = [f_{ij}] \in \mathbb{R}_+^{n \times n}$ has zero diagonal and $f_{ij} = \frac{g_{ij}}{g_{ii}}$ for $i \neq j$

$\gamma_{nls}(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p})^{-1}$

$\Phi_{\mathbf{w},rel}(\gamma) := \sum_{i=1}^n w_i \log \gamma_i, \ \gamma > \mathbf{0}$
obtained by replacing $\log(1+t)$ with smaller $\log t$

Relaxed problem
$\max_{\mathbf{p} \geq \mathbf{0}} \Phi_{\mathbf{w},rel}(\gamma_{nls}) = \max_{\mathbf{p} > \mathbf{0}} \sum_{i=1}^n \mathbf{w}_i \log \frac{p_i}{(F\mathbf{p})_i}$

Relaxation problem

For $\mathbf{z} = (z_1, \ldots, z_n)^\top > \mathbf{0}$ let $\mathbf{z}^{-1} := (z_1^{-1}, \ldots, z_n^{-1})^\top$
$\gamma(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p} + \boldsymbol{\mu})^{-1}$, $\boldsymbol{\mu} = (\frac{\nu_1}{g_{11}}, \ldots, \frac{\nu_n}{g_{nn}})^\top$
$F = [f_{ij}] \in \mathbb{R}_+^{n\times n}$ has zero diagonal and $f_{ij} = \frac{g_{ij}}{g_{ii}}$ for $i \neq j$

$\gamma_{nls}(\mathbf{p}) = \mathbf{p} \circ (F\mathbf{p})^{-1}$

$\Phi_{\mathbf{w},rel}(\gamma) := \sum_{i=1}^n w_i \log \gamma_i, \ \gamma > \mathbf{0}$
obtained by replacing $\log(1+t)$ with smaller $\log t$

Relaxed problem
$\max_{\mathbf{p}\geq\mathbf{0}} \Phi_{\mathbf{w},rel}(\gamma_{nls}) = \max_{\mathbf{p}>\mathbf{0}} \sum_{i=1}^n \mathbf{w}_i \log \frac{p_i}{(F\mathbf{p})_i}$

If $\sum_{j\neq i} w_j > w_i > 0$ for $i = 1, \ldots, n$
relaxed maximal problem can be solved by THM 4.

SIR domain

SIR domain

CLAIM: $\Gamma := \gamma(\mathbb{R}^n_+) := \{\gamma \in \mathbb{R}^n_+,\ \rho(D(\gamma)F) < 1\}$
The inverse map $P : \Gamma \to \mathbb{R}^n_+$ given

$$P(\gamma) = (I - D(\gamma)F)^{-1}(\gamma \circ \boldsymbol{\mu}) = (\sum_{m=0}^{\infty}(D(\gamma)F)^m)(\gamma \circ \boldsymbol{\mu})$$

SIR domain

CLAIM: $\Gamma := \gamma(\mathbb{R}^n_+) := \{\gamma \in \mathbb{R}^n_+,\ \rho(D(\gamma)F) < 1\}$
The inverse map $P : \Gamma \to \mathbb{R}^n_+$ given

$$P(\gamma) = (I - D(\gamma)F)^{-1}(\gamma \circ \boldsymbol{\mu}) = (\sum_{m=0}^{\infty}(D(\gamma)F)^m)(\gamma \circ \boldsymbol{\mu})$$

COR: P increases on Γ: $P(\gamma) < P(\delta)$ if $\gamma \lneq \delta \in \Gamma$.

SIR domain

CLAIM: $\Gamma := \gamma(\mathbb{R}^n_+) := \{\gamma \in \mathbb{R}^n_+,\ \rho(D(\gamma)F) < 1\}$
The inverse map $P : \Gamma \to \mathbb{R}^n_+$ given

$$P(\gamma) = (I - D(\gamma)F)^{-1}(\gamma \circ \boldsymbol{\mu}) = (\sum_{m=0}^{\infty}(D(\gamma)F)^m)(\gamma \circ \boldsymbol{\mu})$$

COR: P increases on Γ: $P(\gamma) < P(\delta)$ if $\gamma \lneq \boldsymbol{\delta} \in \Gamma$.

COR: $\mathbf{p}^\star = (p_1^\star, \ldots, p_n^\star)^\top$ satisfies $p_i^\star = \bar{p}_i$ for some $i = 1, \ldots, n$

SIR domain

CLAIM: $\Gamma := \gamma(\mathbb{R}^n_+) := \{\gamma \in \mathbb{R}^n_+,\ \rho(D(\gamma)F) < 1\}$
The inverse map $P : \Gamma \to \mathbb{R}^n_+$ given

$$P(\gamma) = (I - D(\gamma)F)^{-1}(\gamma \circ \boldsymbol{\mu}) = (\sum_{m=0}^{\infty}(D(\gamma)F)^m)(\gamma \circ \boldsymbol{\mu})$$

COR: P increases on Γ: $P(\gamma) < P(\boldsymbol{\delta})$ if $\gamma \lneq \boldsymbol{\delta} \in \Gamma$.

COR: $\mathbf{p}^\star = (p_1^\star, \dots, p_n^\star)^\top$ satisfies $p_i^\star = \bar{p}_i$ for some $i = 1, \dots, n$

DEF: $[0, p_i] \times \mathbb{R}^{n-1}_+ := \{\mathbf{p} = (p_1, \dots, p_n)^\top \in \mathbb{R}^n_+,\ p_i \le \bar{p}_i\}$,
$\mathbf{e}_i = (\delta_{1i}, \dots, \delta_{ni})^\top$

SIR domain

CLAIM: $\Gamma := \gamma(\mathbb{R}^n_+) := \{\gamma \in \mathbb{R}^n_+,\ \rho(D(\gamma)F) < 1\}$
The inverse map $P : \Gamma \to \mathbb{R}^n_+$ given

$$P(\gamma) = (I - D(\gamma)F)^{-1}(\gamma \circ \boldsymbol{\mu}) = (\sum_{m=0}^{\infty}(D(\gamma)F)^m)(\gamma \circ \boldsymbol{\mu})$$

COR: P increases on Γ: $P(\gamma) < P(\delta)$ if $\gamma \lneq \boldsymbol{\delta} \in \Gamma$.

COR: $\mathbf{p}^\star = (p_1^\star, \ldots, p_n^\star)^\top$ satisfies $p_i^\star = \bar{p}_i$ for some $i = 1, \ldots, n$

DEF: $[0, p_i] \times \mathbb{R}^{n-1}_+ := \{\mathbf{p} = (p_1, \ldots, p_n)^\top \in \mathbb{R}^n_+,\ p_i \le \bar{p}_i\}$,
$\mathbf{e}_i = (\delta_{1i}, \ldots, \delta_{ni})^\top$

THM 5: $\gamma([0, p_i] \times \mathbb{R}^{n-1}_+) = \{\gamma \in \mathbb{R}^n_+,\ \rho(D(\gamma)(F + \frac{1}{\bar{p}_i}\boldsymbol{\mu}\mathbf{e}_i^\top)) \le 1\}$

SIR domain

CLAIM: $\Gamma := \gamma(\mathbb{R}^n_+) := \{\gamma \in \mathbb{R}^n_+, \ \rho(D(\gamma)F) < 1\}$
The inverse map $P : \Gamma \to \mathbb{R}^n_+$ given

$$P(\gamma) = (I - D(\gamma)F)^{-1}(\gamma \circ \boldsymbol{\mu}) = (\sum_{m=0}^{\infty}(D(\gamma)F)^m)(\gamma \circ \boldsymbol{\mu})$$

COR: P increases on Γ: $P(\gamma) < P(\boldsymbol{\delta})$ if $\gamma \lneq \boldsymbol{\delta} \in \Gamma$.

COR: $\mathbf{p}^\star = (p_1^\star, \ldots, p_n^\star)^\top$ satisfies $p_i^\star = \bar{p}_i$ for some $i = 1, \ldots, n$

DEF: $[0, p_i] \times \mathbb{R}^{n-1}_+ := \{\mathbf{p} = (p_1, \ldots, p_n)^\top \in \mathbb{R}^n_+, \ p_i \leq \bar{p}_i\}$,
$\mathbf{e}_i = (\delta_{1i}, \ldots, \delta_{ni})^\top$

THM 5: $\gamma([0, p_i] \times \mathbb{R}^{n-1}_+) = \{\gamma \in \mathbb{R}^n_+, \ \rho(D(\gamma)(F + \frac{1}{\bar{p}_i}\boldsymbol{\mu}\mathbf{e}_i^\top)) \leq 1\}$

COR $\gamma([\mathbf{0}, \mathbf{p}]) = \{\gamma \in \mathbb{R}^n_+, \ \rho(D(\gamma)(F + \frac{1}{\bar{p}_i}\boldsymbol{\mu}\mathbf{e}_i^\top)) \leq 1, i = 1, \ldots, n\}$

Restatement of the maximal problem

Restatement of the maximal problem

$\mathbf{0} < \gamma = e^{\log \gamma}$. New variable $\mathbf{x} = \log \gamma$

Restatement of the maximal problem

$\mathbf{0} < \gamma = e^{\log \gamma}$. New variable $\mathbf{x} = \log \gamma$
Hence $\log \gamma([\mathbf{0}, \mathbf{p}])$ is the closed unbounded closed set $\mathcal{D} \subset \mathbb{R}^n$:

$$h_i(\mathbf{x}) := \log \rho(\mathrm{diag}(e^{\mathbf{x}})(F + \frac{1}{\bar{p}_i}\boldsymbol{\mu}\mathbf{e}_i^\top)) \leq 0, \quad i = 1, \ldots, n$$

Restatement of the maximal problem

$\mathbf{0} < \gamma = e^{\log\gamma}$. New variable $\mathbf{x} = \log\gamma$
Hence $\log\gamma([\mathbf{0}, \mathbf{p}])$ is the closed unbounded closed set $\mathcal{D} \subset \mathbb{R}^n$:

$$h_i(\mathbf{x}) := \log\rho(\mathrm{diag}(e^{\mathbf{x}})(F + \frac{1}{\bar{p}_i}\boldsymbol{\mu}\mathbf{e}_i^\top)) \le 0, \quad i = 1, \ldots, n$$

Since $h_i(\mathbf{x})$ is convex, $\mathcal{D}$ convex

Restatement of the maximal problem

$\mathbf{0} < \gamma = e^{\log \gamma}$. New variable $\mathbf{x} = \log \gamma$
Hence $\log \gamma([\mathbf{0}, \mathbf{p}])$ is the closed unbounded closed set $\mathcal{D} \subset \mathbb{R}^n$:

$$h_i(\mathbf{x}) := \log \rho(\mathrm{diag}(e^{\mathbf{x}})(F + \frac{1}{\bar{p}_i}\boldsymbol{\mu}\mathbf{e}_i^\top)) \le 0, \quad i = 1, \ldots, n$$

Since $h_i(\mathbf{x})$ is convex, $\mathcal{D}$ convex
Since $\log(1 + e^t)$ convex, the equivalent maximal problem

$$\max_{\mathbf{x} \in \mathcal{D}} \Phi_{\mathbf{w}}(e^{\mathbf{x}}) = \max_{\mathbf{x}, h_i(\mathbf{x}) \le 0, i=1,\ldots,n} \sum_{j=1}^{n} \log(1 + e^{x_j})$$

Restatement of the maximal problem

$\mathbf{0} < \gamma = e^{\log \gamma}$. New variable $\mathbf{x} = \log \gamma$
Hence $\log \gamma([\mathbf{0}, \mathbf{p}])$ is the closed unbounded closed set $\mathcal{D} \subset \mathbb{R}^n$:

$$h_i(\mathbf{x}) := \log \rho(\mathrm{diag}(e^{\mathbf{x}})(F + \frac{1}{\bar{p}_i}\boldsymbol{\mu}\mathbf{e}_i^\top)) \le 0, \quad i = 1, \ldots, n$$

Since $h_i(\mathbf{x})$ is convex, $\mathcal{D}$ convex
Since $\log(1 + e^t)$ convex, the equivalent maximal problem

$$\max_{\mathbf{x} \in \mathcal{D}} \Phi_{\mathbf{w}}(e^{\mathbf{x}}) = \max_{\mathbf{x}, h_i(\mathbf{x}) \le 0, i=1,\ldots,n} \sum_{j=1}^{n} \log(1 + e^{x_j})$$

maximization of convex function on closed unbounded convex set

Approximation methods-I

Approximation 1:
For $K \gg 1$ $\mathcal{D}_K := \{\mathbf{x} \in \mathcal{D},\ \mathbf{x} \geq -K\mathbf{1} = -K(1,\ldots,1)^\top\}$
consider $\max_{\mathbf{x}\in\mathcal{D}_K} \Phi_{\mathbf{w}}$

Approximation methods-I

Approximation 1:
For $K \gg 1$ $\mathcal{D}_K := \{\mathbf{x} \in \mathcal{D},\ \mathbf{x} \geq -K\mathbf{1} = -K(1,\ldots,1)^\top\}$
consider $\max_{\mathbf{x}\in\mathcal{D}_K} \Phi_{\mathbf{w}}$

Approximation 2:
Choose a few boundary points $\xi_1,\ldots,\xi_N \in \mathcal{D}$ s.t.
$h_j(\xi_k) = 0$ for $j \in \mathcal{A}_k \subset \{1,\ldots,n\}$ and $k = 1,\ldots,N$.

Approximation methods-I

Approximation 1:
For $K \gg 1$ $\mathcal{D}_K := \{\mathbf{x} \in \mathcal{D},\ \mathbf{x} \geq -K\mathbf{1} = -K(1,\ldots,1)^\top\}$
consider $\max_{\mathbf{x}\in\mathcal{D}_K} \Phi_{\mathbf{w}}$

Approximation 2:
Choose a few boundary points $\xi_1,\ldots,\xi_N \in \mathcal{D}$ s.t.
$h_j(\xi_k) = 0$ for $j \in \mathcal{A}_k \subset \{1,\ldots,n\}$ and $k = 1,\ldots,N$.
At each ξ_k one has $\#\mathcal{A}_k$ supporting hyperplanes $H_{j,k}, j \in \mathcal{A}_k$

Approximation methods-I

Approximation 1:
For $K \gg 1$ $\mathcal{D}_K := \{\mathbf{x} \in \mathcal{D},\ \mathbf{x} \ge -K\mathbf{1} = -K(1,\ldots,1)^\top\}$
consider $\max_{\mathbf{x}\in\mathcal{D}_K} \Phi_{\mathbf{w}}$

Approximation 2:
Choose a few boundary points $\xi_1,\ldots,\xi_N \in \mathcal{D}$ s.t.
$h_j(\xi_k) = 0$ for $j \in \mathcal{A}_k \subset \{1,\ldots,n\}$ and $k = 1,\ldots,N$.
At each ξ_k one has $\#\mathcal{A}_k$ supporting hyperplanes $H_{j,k}, j \in \mathcal{A}_k$

The supporting hyperplane of $h_j(\mathbf{x})$ at ξ_k is $H_{j,k}(\mathbf{x}) \le H_{j,k}(\xi_k)$
$H_{j,k}(\mathbf{x}) = \mathbf{w}_{j,k}^\top \mathbf{x},\ \mathbf{w}_{j,k} = \mathbf{x}(D(e^{\xi_k})(F + \frac{1}{\bar{p}_j}\mu\mathbf{e}_j^\top)) \circ \mathbf{y}(D(e^{\xi_k})(F + \frac{1}{\bar{p}_j}\mu\mathbf{e}_j^\top))$

$$\mathcal{D}(\xi_1,\ldots,\xi_N,K) = \{\mathbf{x} \in \mathbb{R}^n, H_{j,k}(\mathbf{x}) \le H_{j,k}(\xi_k), j \in \mathcal{A}_k, k \in \langle N\rangle, \xi \ge -K\mathbf{1}\}$$

$$\mathcal{D}_K \subset \mathcal{D}(\xi_1,\ldots,\xi_N,K)$$

Approximation methods-I

Approximation 1:
For $K \gg 1$ $\mathcal{D}_K := \{\mathbf{x} \in \mathcal{D},\ \mathbf{x} \geq -K\mathbf{1} = -K(1,\ldots,1)^\top\}$
consider $\max_{\mathbf{x}\in\mathcal{D}_K} \Phi_{\mathbf{w}}$

Approximation 2:
Choose a few boundary points $\xi_1,\ldots,\xi_N \in \mathcal{D}$ s.t.
$h_j(\xi_k) = 0$ for $j \in \mathcal{A}_k \subset \{1,\ldots,n\}$ and $k = 1,\ldots,N$.
At each ξ_k one has $\#\mathcal{A}_k$ supporting hyperplanes $H_{j,k}, j \in \mathcal{A}_k$

The supporting hyperplane of $h_j(\mathbf{x})$ at ξ_k is $H_{j,k}(\mathbf{x}) \leq H_{j,k}(\xi_k)$
$H_{j,k}(\mathbf{x}) = \mathbf{w}_{j,k}^\top \mathbf{x},\ \mathbf{w}_{j,k} = \mathbf{x}(D(e^{\xi_k})(F + \frac{1}{\bar{p}_j}\mu\mathbf{e}_j^\top)) \circ \mathbf{y}(D(e^{\xi_k})(F + \frac{1}{\bar{p}_j}\mu\mathbf{e}_j^\top))$

$\mathcal{D}(\xi_1,\ldots,\xi_N,K) = \{\mathbf{x} \in \mathbb{R}^n, H_{j,k}(\mathbf{x}) \leq H_{j,k}(\xi_k), j \in \mathcal{A}_k, k \in \langle N\rangle, \xi \geq -K\mathbf{1}\}$

$\mathcal{D}_K \subset \mathcal{D}(\xi_1,\ldots,\xi_N,K)$

$\max_{\mathbf{x}\in\mathcal{D}(\xi_1,\ldots,\xi_N,K)} \Phi_{\mathbf{w}}(e^{\mathbf{x}}) \geq \max_{\mathbf{x}\in\mathcal{D}_K} \Phi_{\mathbf{w}}(e^{\mathbf{x}})$

Approximation methods-II

Approximation 3:

$$\max_{\mathbf{x}\in\mathcal{D}(\xi_1,\ldots,\xi_N,K)} \Phi_{\mathbf{w},rel}(e^{\mathbf{x}}) = \max_{\mathbf{x}\in\mathcal{D}(\xi_1,\ldots,\xi_N,K)} \mathbf{w}^\top \mathbf{x}$$

Use Simplex Method or Ellipsoid Algorithm

Approximation methods-II

Approximation 3:

$$\max_{\mathbf{x}\in\mathcal{D}(\xi_1,\ldots,\xi_N,K)} \Phi_{\mathbf{w},rel}(e^{\mathbf{x}}) = \max_{\mathbf{x}\in\mathcal{D}(\xi_1,\ldots,\xi_N,K)} \mathbf{w}^\top \mathbf{x}$$

Use Simplex Method or Ellipsoid Algorithm

Choice of $\xi_1,\ldots,\xi_N$:

Pick a finite number $\mathbf{0} < \mathbf{p}_1,\ldots,\mathbf{p}_N \in [\mathbf{0},\bar{\mathbf{p}}] = [0,\bar{p}_1]\times\ldots[0,\bar{p}_n]$ boundary points

Approximation methods-II

Approximation 3:

$$\max_{\mathbf{x}\in\mathcal{D}(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_N,K)} \Phi_{\mathbf{w},rel}(e^{\mathbf{x}}) = \max_{\mathbf{x}\in\mathcal{D}(\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_N,K)} \mathbf{w}^\top\mathbf{x}$$

Use Simplex Method or Ellipsoid Algorithm

Choice of $\xi_1, \ldots, \xi_N$:

Pick a finite number $\mathbf{0} < \mathbf{p}_1, \ldots, \mathbf{p}_N \in [\mathbf{0}, \bar{\mathbf{p}}] = [0, \bar{p}_1] \times \ldots [0, \bar{p}_n]$ boundary points
E.g., divide $[\mathbf{0}, \mathbf{p}]$ by a mesh, and choose all boundary points with positive coordinates

$\xi_k = \gamma(\mathbf{p}_k)$ and $\mathcal{A}_k$ all j s.t. $p_{j,k} = \bar{p}_j$

Direct methods

Study $\max_{\mathbf{0}\le\mathbf{p}\le\bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\mathbf{p}) = \Phi_{\mathbf{w}}(\mathbf{p}^\star)$

If $w_i = 0$ then $p_i^\star = 0$.

Direct methods

Study $\max_{\mathbf{0}\le\mathbf{p}\le\bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\mathbf{p}) = \Phi_{\mathbf{w}}(\mathbf{p}^\star)$

If $w_i = 0$ then $p_i^\star = 0$.

Assumption $\mathbf{0} < \mathbf{w} \in \Pi_n$

Direct methods

Study $\max_{\mathbf{0}\le\mathbf{p}\le\bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\mathbf{p}) = \Phi_{\mathbf{w}}(\mathbf{p}^\star)$

If $w_i = 0$ then $p_i^\star = 0$.

Assumption $\mathbf{0} < \mathbf{w} \in \Pi_n$

Local minimum conditions at $\mathbf{0} \ne p^\star \in \partial[\mathbf{0}, \bar{\mathbf{p}}]$

Direct methods

Study $\max_{\mathbf{0}\le\mathbf{p}\le\bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\mathbf{p}) = \Phi_{\mathbf{w}}(\mathbf{p}^\star)$

If $w_i = 0$ then $p_i^\star = 0$.

Assumption $\mathbf{0} < \mathbf{w} \in \Pi_n$

Local minimum conditions at $\mathbf{0} \ne p^\star \in \partial[\mathbf{0}, \bar{\mathbf{p}}]$

1. $\partial_i \Phi_{\mathbf{w}}(\mathbf{p}^\star) = 0$ if $0 < p_i^\star < \bar{p}_i$

Direct methods

Study $\max_{\mathbf{0}\le\mathbf{p}\le\bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\mathbf{p}) = \Phi_{\mathbf{w}}(\mathbf{p}^\star)$

If $w_i = 0$ then $p_i^\star = 0$.

Assumption $\mathbf{0} < \mathbf{w} \in \Pi_n$

Local minimum conditions at $\mathbf{0} \neq p^\star \in \partial[\mathbf{0}, \bar{\mathbf{p}}]$

1. $\partial_i \Phi_{\mathbf{w}}(\mathbf{p}^\star) = 0$ if $0 < p_i^\star < \bar{p}_i$

2. $\partial_i \Phi_{\mathbf{w}}(\mathbf{p}^\star) \ge 0$ if $p_i^\star = \bar{p}_i$

Direct methods

Study $\max_{\mathbf{0} \le \mathbf{p} \le \bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\mathbf{p}) = \Phi_{\mathbf{w}}(\mathbf{p}^\star)$

If $w_i = 0$ then $p_i^\star = 0$.

Assumption $\mathbf{0} < \mathbf{w} \in \Pi_n$

Local minimum conditions at $\mathbf{0} \ne p^\star \in \partial[\mathbf{0}, \bar{\mathbf{p}}]$

1. $\partial_i \Phi_{\mathbf{w}}(\mathbf{p}^\star) = 0$ if $0 < p_i^\star < \bar{p}_i$
2. $\partial_i \Phi_{\mathbf{w}}(\mathbf{p}^\star) \ge 0$ if $p_i^\star = \bar{p}_i$
3. $\partial_i \Phi_{\mathbf{w}}(\mathbf{p}^\star) \le 0$ if $p_i^\star = 0$

Direct methods

Study $\max_{\mathbf{0}\le\mathbf{p}\le\bar{\mathbf{p}}} \Phi_{\mathbf{w}}(\mathbf{p}) = \Phi_{\mathbf{w}}(\mathbf{p}^\star)$

If $w_i = 0$ then $p_i^\star = 0$.

Assumption $\mathbf{0} < \mathbf{w} \in \Pi_n$

Local minimum conditions at $\mathbf{0} \ne p^\star \in \partial[\mathbf{0}, \bar{\mathbf{p}}]$

1. $\partial_i \Phi_{\mathbf{w}}(\mathbf{p}^\star) = 0$ if $0 < p_i^\star < \bar{p}_i$

2. $\partial_i \Phi_{\mathbf{w}}(\mathbf{p}^\star) \ge 0$ if $p_i^\star = \bar{p}_i$

3. $\partial_i \Phi_{\mathbf{w}}(\mathbf{p}^\star) \le 0$ if $p_i^\star = 0$

Apply gradient methods and their variations

References 1

- R.A. Brualdi, S.V. Parter and H. Schneider, The diagonal equivalence of a nonnegative matrix to a stochastic matrix, *J. Math. Anal. Appl.* 16 (1966), 31–50.
- J.E. Cohen, Random evolutions and the spectral radius of a non-negative matrix, *Mat. Proc. Camb. Phil. Soc* 86 (1979), 345-350.
- M.D. Donsker and S.R.S. Varadhan, On a variational formula for the principal eigenvalue for operators with maximum principle, *Proc. Nat. Acad. Sci. U.S.A.* 72 (1975), 780–783.
- S. Friedland, Convex spectral functions, *Linear Multilin. Algebra* 9 (1981), 299-316.
- S. Friedland and S. Karlin, Some inequalities for the spectral radius of non-negative matrices and applications, *Duke Mathematical Journal* 42 (3), 459-490, 1975.

References 2

- S. Friedland and C.W. Tan, Maximizing sum rates in Gaussian interference-limited channels, arXiv:0806.2860
- J. F. C. Kingman, A convexity property of positive matrices, *Quart. J. Math. Oxford Ser.* 12 (2), 283-284, 1961.
- R. A. Sinkhorn, A relationship between arbitrary positive matrices and doubly stochastic matrices, *Ann. Math. Statist.* 35 (1964), 876–879.
- R. Sinkhorn and P. Knopp, Concerning nonnegative matrices and doubly stochastic matrices, *Pacific J. Math.* 21 (1967), 343–348.