Block designs, spanning trees and resistance in electrical networks

R. A. Bailey

Queen Mary
University of London

r.a.bailey@qmul.ac.uk

British Combinatorial Conference
July 2007

Abstract

The last decade has seen a huge number of microarray experiments performed. If we ignore some complications, the designs for these experiments are just block designs with block size two, so they can also be regarded as graphs, possibly with multiple edges.
Statisticians usually rate block designs by (at least) two criteria. When the block size is two, a design is D-optimal if it has the maximum number of spanning trees, and it is A-optimal if it has the minimum total of pairwise resistances when there is a unit resistance on each edge (block). Experience with block designs in other situations suggests that these two criteria agree closely at the top end. However, microarray experiments are usually done with such a small number b of blocks, relative to the number t of treatments, that the two criteria give opposite ranks.
I shall show that, if b is too small relative to t, this happens for all sufficiently large t.

Microarray experiments

The last decade has seen a huge number of microarray experiments performed in genomics.

If we ignore some complications,
the designs for these experiments are just block designs with block size two,
so they can also be regarded as graphs, possibly with multiple edges.

Microarray experiments

The last decade has seen a huge number of microarray experiments performed in genomics.

If we ignore some complications,
the designs for these experiments are just block designs with block size two,
so they can also be regarded as graphs, possibly with multiple edges.

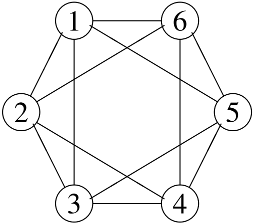
$$v \text{ treatments} \longrightarrow v \text{ vertices}$$
$$b \text{ blocks} \longrightarrow b \text{ edges}$$

Designs for 6 treatments (6 vertices)

12 blocks (edges)

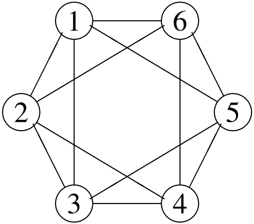
Designs for 6 treatments (6 vertices)

12 blocks (edges)



Designs for 6 treatments (6 vertices)

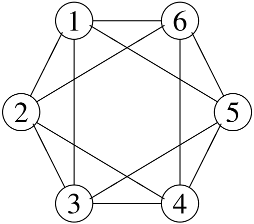
12 blocks (edges)



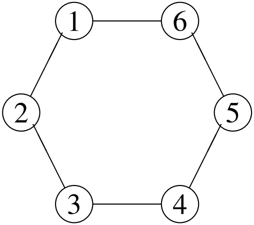
6 blocks (edges)

Designs for 6 treatments (6 vertices)

12 blocks (edges)

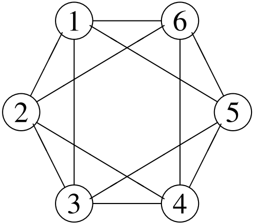


6 blocks (edges)

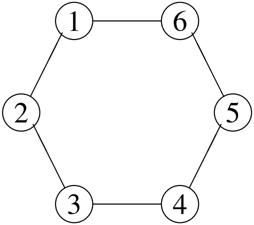


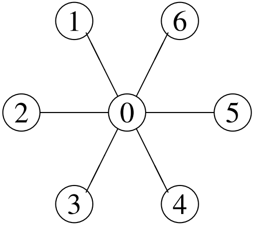
Designs for 6 treatments (6 vertices)

12 blocks (edges)



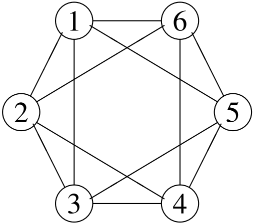
6 blocks (edges)

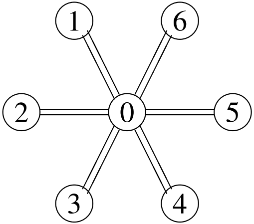




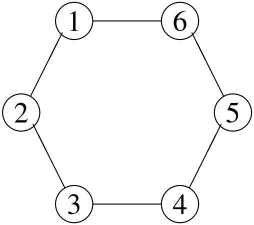
Designs for 6 treatments (6 vertices)

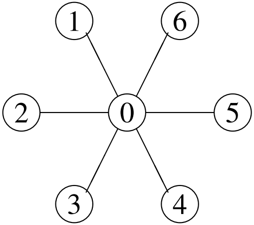
12 blocks (edges)



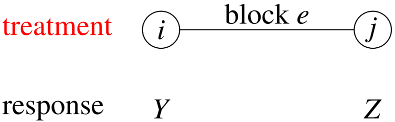


6 blocks (edges)





Estimation and variance



Estimation and variance

treatment

response $\quad Y \qquad Z$

$$\text{Assume that } Y = \tau_i + \beta_e + \text{noise}$$
$$Z = \tau_j + \beta_e + \text{other noise}$$

Estimation and variance

treatment 

response Y Z

$$
\begin{aligned}
\text{Assume that } Y &= \tau_i + \beta_e + \text{noise} \\
Z &= \tau_j + \beta_e + \text{other noise}
\end{aligned}
$$

Then $Y - Z$ is an estimator for $\tau_i - \tau_j$ with variance 1.

Estimation and variance

treatment 

response Y Z

$$\begin{aligned}\text{Assume that } Y &= \tau_i + \beta_e + \text{noise}\\ Z &= \tau_j + \beta_e + \text{other noise}\end{aligned}$$

Then $Y - Z$ is an estimator for $\tau_i - \tau_j$ with variance 1.

Put $V_{ij} =$ variance of the estimator for $\tau_i - \tau_j$ using the whole graph.

We want all the V_{ij} to be small.

Optimality

Put $A =$ adjacency matrix of the graph
(so $a_{ij} =$ number of blocks containing i and j);

Optimality

Put $A =$ adjacency matrix of the graph
(so $a_{ij} =$ number of blocks containing i and j);
and $R =$ diagonal matrix of replications
(so $r_{ii} =$ valency of vertex i).

Optimality

Put $A =$ adjacency matrix of the graph
(so $a_{ij} =$ number of blocks containing i and j);
and $R =$ diagonal matrix of replications
(so $r_{ii} =$ valency of vertex i).
The information matrix is $L = R - (A+R)/2 =$ Laplacian matrix $\div 2$.

Optimality

Put $A =$ adjacency matrix of the graph
(so $a_{ij} =$ number of blocks containing i and j);
and $R =$ diagonal matrix of replications
(so $r_{ii} =$ valency of vertex i).
The information matrix is $L = R - (A+R)/2 =$ Laplacian matrix $\div 2$.
It has a "trivial" eigenvalue 0, corresponding to the all-1 eigenvector.

Optimality

Put $A =$ adjacency matrix of the graph
(so $a_{ij} =$ number of blocks containing i and j);
and $R =$ diagonal matrix of replications
(so $r_{ii} =$ valency of vertex i).
The information matrix is $L = R - (A+R)/2 =$ Laplacian matrix $\div 2$.
It has a "trivial" eigenvalue 0, corresponding to the all-1 eigenvector.

The design is called

- A-optimal if it maximizes the harmonic mean of the non-trivial eigenvalues;

over all block designs with block size two and the given v and b.

Optimality

Put $A =$ adjacency matrix of the graph
(so $a_{ij} =$ number of blocks containing i and j);
and $R =$ diagonal matrix of replications
(so $r_{ii} =$ valency of vertex i).
The information matrix is $L = R - (A+R)/2 =$ Laplacian matrix $\div\, 2$.
It has a "trivial" eigenvalue 0, corresponding to the all-1 eigenvector.

The design is called

- A-optimal if it maximizes the harmonic mean of the non-trivial eigenvalues;
- D-optimal if it maximizes the geometric mean of the non-trivial eigenvalues

over all block designs with block size two and the given v and b.

Optimality

Put $A =$ adjacency matrix of the graph
(so $a_{ij} =$ number of blocks containing i and j);
and $R =$ diagonal matrix of replications
(so $r_{ii} =$ valency of vertex i).
The information matrix is $L = R - (A+R)/2 =$ Laplacian matrix $\div 2$.
It has a "trivial" eigenvalue 0, corresponding to the all-1 eigenvector.

The design is called

- A-optimal if it maximizes the harmonic mean of the non-trivial eigenvalues;
  —equivalently, it minimizes the average of the variances V_{ij}
- D-optimal if it maximizes the geometric mean of the non-trivial eigenvalues

over all block designs with block size two and the given v and b.

Optimality

Put $A =$ adjacency matrix of the graph
(so $a_{ij} =$ number of blocks containing i and j);
and $R =$ diagonal matrix of replications
(so $r_{ii} =$ valency of vertex i).
The information matrix is $L = R - (A+R)/2 =$ Laplacian matrix $\div\, 2$.
It has a "trivial" eigenvalue 0, corresponding to the all-1 eigenvector.

The design is called

- A-optimal if it maximizes the harmonic mean of the non-trivial eigenvalues;
  —equivalently, it minimizes the average of the variances V_{ij}
- D-optimal if it maximizes the geometric mean of the non-trivial eigenvalues
  —equivalently, it minimizes the volume of the confidence ellipsoid for $(\tau_1, \ldots, \tau_v)$

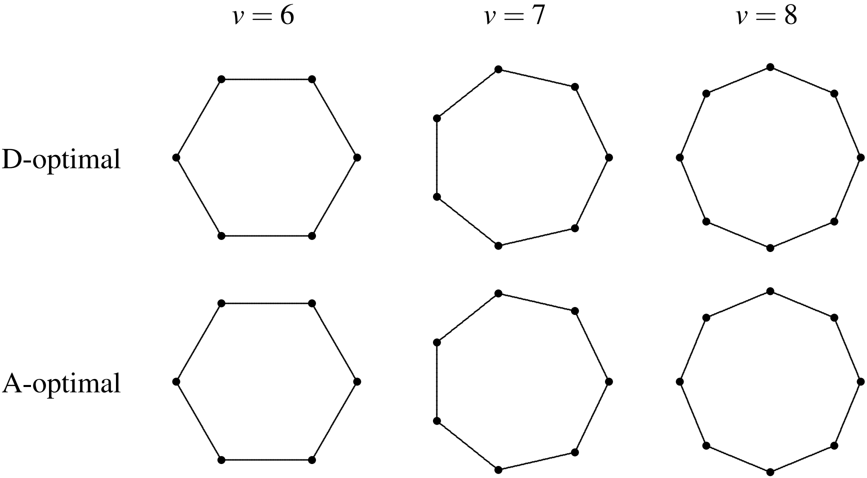
over all block designs with block size two and the given v and b.

What happens when $b = v$?

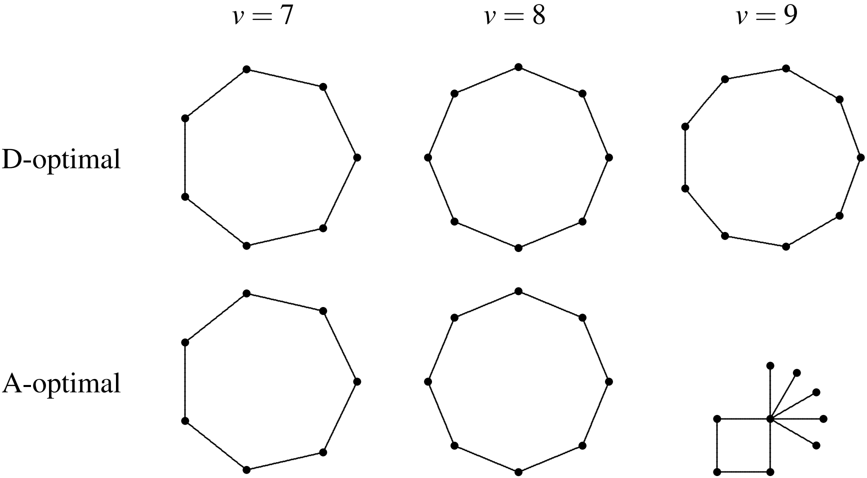
Computer investigation by

- Jones and Eccleston (1980)
- Kerr and Churchill (2001)
- Wit, Nobile and Khanin (2005)
- Ceraudo (2005).

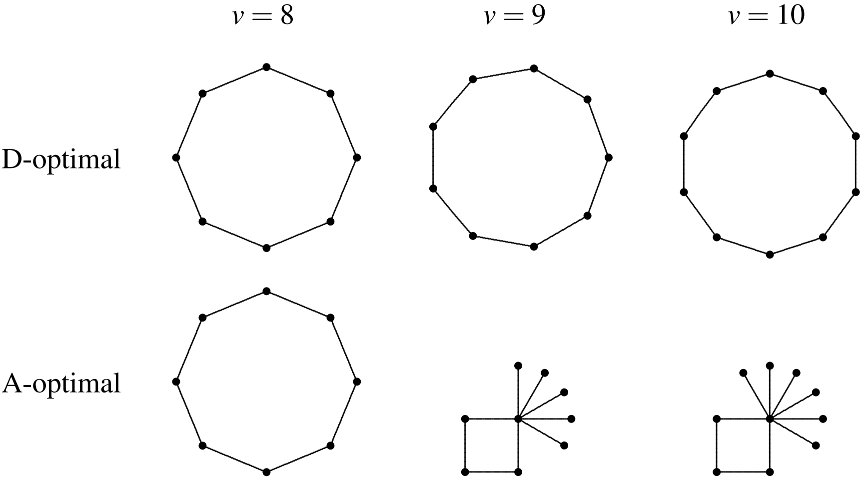
Optimal designs when $b = v$



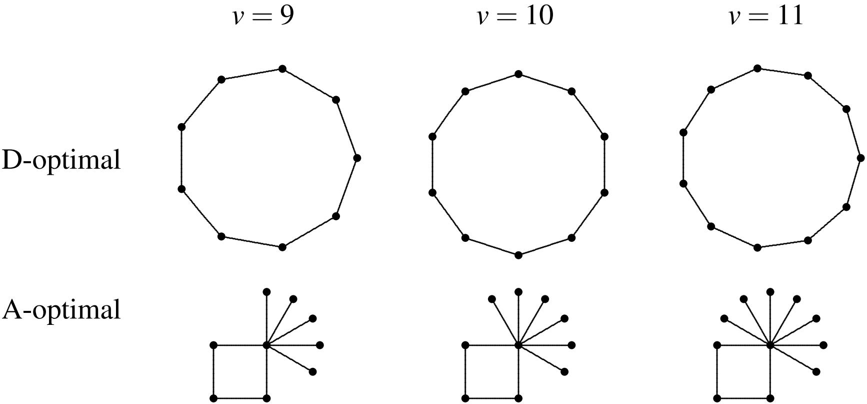
Optimal designs when $b = v$



Optimal designs when $b = v$



Optimal designs when $b = v$



D-optimality

Cheng (1978), after Gaffke (1978), after Kirchhoff (1847):

$$\text{product of non-trivial eigenvalues of } L = \frac{v \times \text{number of spanning trees}}{2^{v-1}}$$

D-optimality

Cheng (1978), after Gaffke (1978), after Kirchhoff (1847):

$$\text{product of non-trivial eigenvalues of } L = \frac{v \times \text{number of spanning trees}}{2^{v-1}}$$

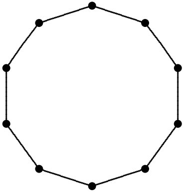
number of spanning trees $=$
number of ways of removing $b-v+1$ edges without disconnecting the graph, (which is easy to calculate by hand when $b-v$ is small)

D-optimality

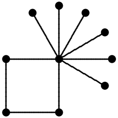
Cheng (1978), after Gaffke (1978), after Kirchhoff (1847):

$$\text{product of non-trivial eigenvalues of } L = \frac{v \times \text{number of spanning trees}}{2^{v-1}}$$

number of spanning trees $=$
number of ways of removing $b - v + 1$ edges without disconnecting the graph, (which is easy to calculate by hand when $b - v$ is small)



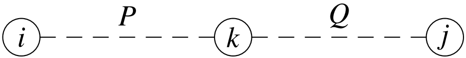
10 spanning trees



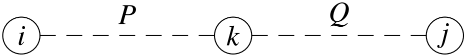
4 spanning trees

The loop design is uniquely D-optimal when $b = v$.

A-optimality: series



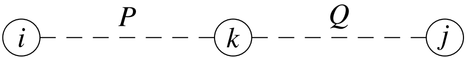
A-optimality: series



In our block design,

$$\begin{aligned} P &= \text{estimator for } \tau_i - \tau_k \text{ with variance } d \\ Q &= \text{estimator for } \tau_k - \tau_j \text{ with variance } e \end{aligned}$$

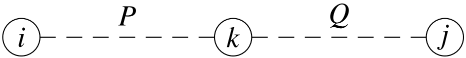
A-optimality: series



In our block design,

$$
\begin{aligned}
P &= \text{estimator for } \tau_i - \tau_k \text{ with variance } d \\
Q &= \text{estimator for } \tau_k - \tau_j \text{ with variance } e \\
P + Q &= \text{estimator for } \tau_i - \tau_j \text{ with variance } d + e
\end{aligned}
$$

A-optimality: series



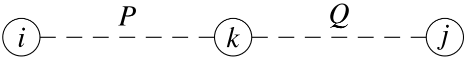
In our block design,

$$\begin{aligned} P &= \text{estimator for } \tau_i - \tau_k \text{ with variance } d \\ Q &= \text{estimator for } \tau_k - \tau_j \text{ with variance } e \\ P+Q &= \text{estimator for } \tau_i - \tau_j \text{ with variance } d+e \end{aligned}$$

In an electrical network,

resistance in network P between i and k	is	d
resistance in network Q between k and j	is	e

A-optimality: series



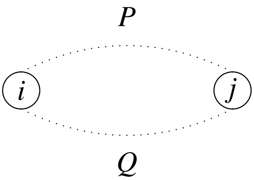
In our block design,

$$
\begin{aligned}
P &= \text{estimator for } \tau_i - \tau_k \text{ with variance } d \\
Q &= \text{estimator for } \tau_k - \tau_j \text{ with variance } e \\
P + Q &= \text{estimator for } \tau_i - \tau_j \text{ with variance } d + e
\end{aligned}
$$

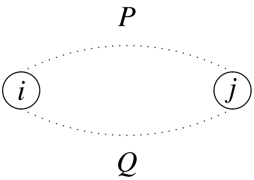
In an electrical network,

resistance in network P between i and k	is	d
resistance in network Q between k and j	is	e
resistance in network 'P in series with Q' between i and j	is	$d + e$

A-optimality: parallel



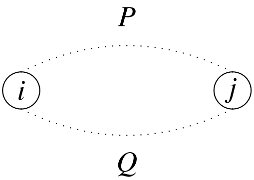
A-optimality: parallel



In our block design,

$$P = \text{estimator for } \tau_i - \tau_j \text{ with variance } d$$
$$Q = \text{estimator for } \tau_i - \tau_j \text{ with variance } e$$

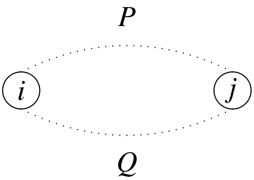
A-optimality: parallel



In our block design,

$$
\begin{aligned}
P &= \text{estimator for } \tau_i - \tau_j \text{ with variance } d \\
Q &= \text{estimator for } \tau_i - \tau_j \text{ with variance } e \\
\frac{eP + dQ}{d + e} &= \text{best estimator for } \tau_i - \tau_j \text{ with variance } \frac{1}{\frac{1}{d} + \frac{1}{e}}
\end{aligned}
$$

A-optimality: parallel



In our block design,

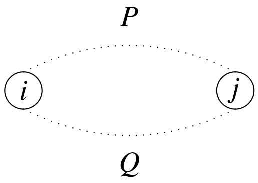
$$\begin{aligned} P &= \text{estimator for } \tau_i - \tau_j \text{ with variance } d \\ Q &= \text{estimator for } \tau_i - \tau_j \text{ with variance } e \\ \frac{eP + dQ}{d+e} &= \text{best estimator for } \tau_i - \tau_j \text{ with variance } \frac{1}{\frac{1}{d} + \frac{1}{e}} \end{aligned}$$

In an electrical network,

resistance in network P between i and j is d

resistance in network Q between i and j is e

A-optimality: parallel



In our block design,

$$\begin{aligned} P &= \text{estimator for } \tau_i - \tau_j \text{ with variance } d \\ Q &= \text{estimator for } \tau_i - \tau_j \text{ with variance } e \\ \frac{eP + dQ}{d+e} &= \text{best estimator for } \tau_i - \tau_j \text{ with variance } \frac{1}{\frac{1}{d} + \frac{1}{e}} \end{aligned}$$

In an electrical network,

resistance in network P between i and j is d

resistance in network Q between i and j is e

resistance in network 'P in parallel with Q' between i and j is $\dfrac{1}{\frac{1}{d} + \frac{1}{e}}$

Eureka!

Variance is the same as resistance!

Block design $\longrightarrow$ (multi-)graph

$\longrightarrow$ electrical network with resistance 1 in each edge

Use Ohm's Law and Kirchhoff's Laws to calculate each variance V_{ij} algebraically.

Results about A-optimality

Theorem

If an A-optimal design has leaves (vertices of valency 1*), then they are all attached to the same vertex.*

Results about A-optimality

Theorem

If an A-optimal design has leaves (vertices of valency 1), then they are all attached to the same vertex.

Theorem

The A-optimal design with v treatments (vertices) and v blocks (edges) is

- *the loop design (single circuit) if* $v \leq 8$;
- *a star glued to a quadrilateral if* $9 \leq v \leq 11$;
- *a star glued to a quadrilateral or a triangle if* $v = 12$;
- *a star glued to a triangle if* $v \geq 13$.

Results about A-optimality

Theorem

If an A-optimal design has leaves (vertices of valency 1*), then they are all attached to the same vertex.*

Theorem

The A-optimal design with v *treatments (vertices) and* v *blocks (edges) is*

- *the loop design (single circuit) if* $v \leq 8$*;*
- *a star glued to a quadrilateral if* $9 \leq v \leq 11$*;*
- *a star glued to a quadrilateral or a triangle if* $v = 12$*;*
- *a star glued to a triangle if* $v \geq 13$*.*

Theorem

The A-optimal design with v *treatments (vertices) and* $v+1$ *blocks (edges) is …*

Threshold

Bad news theorem

Given any fixed value of $b - v$*, there is a threshold* T *such that when* $v \geq T$ *the A- and D-optimality criteria conflict.*

Threshold

Bad news theorem

Given any fixed value of $b-v$, there is a threshold T such that when $v \geq T$ the A- and D-optimality criteria conflict.

Fix $b-v=c$.

$$\text{Average valency } = \frac{2b}{v} = \frac{2(v+c)}{v} = 2 + \frac{2c}{v} < 3 \text{ for large } v.$$

Threshold

Bad news theorem

Given any fixed value of $b-v$, there is a threshold T such that when $v \geq T$ the A- and D-optimality criteria conflict.

Fix $b-v=c$.

$$\text{Average valency } = \frac{2b}{v} = \frac{2(v+c)}{v} = 2 + \frac{2c}{v} < 3 \text{ for large } v.$$

For large v, our graph G consists of

- a graph G_0 with all valencies at least 3;
- some vertices inserted into edges of G_0;
- some leaves attached to one (or more) vertices of G_0.

Threshold

Bad news theorem

Given any fixed value of $b - v$, there is a threshold T such that when $v \geq T$ the A- and D-optimality criteria conflict.

Fix $b - v = c$.

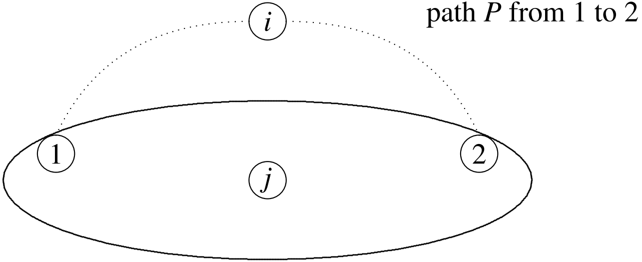
$$\text{Average valency } = \frac{2b}{v} = \frac{2(v+c)}{v} = 2 + \frac{2c}{v} < 3 \text{ for large } v.$$

For large v, our graph G consists of

- a graph G_0 with all valencies at least 3;
- some vertices inserted into edges of G_0;
- some leaves attached to one (or more) vertices of G_0.

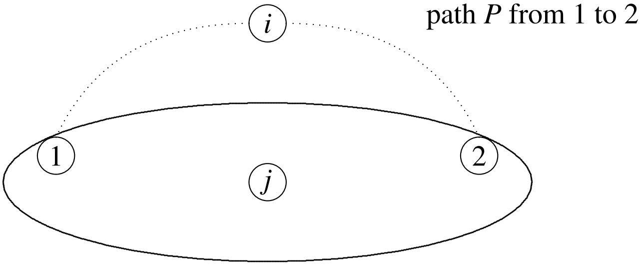
D-optimal designs do not have leaves.

Proof outline



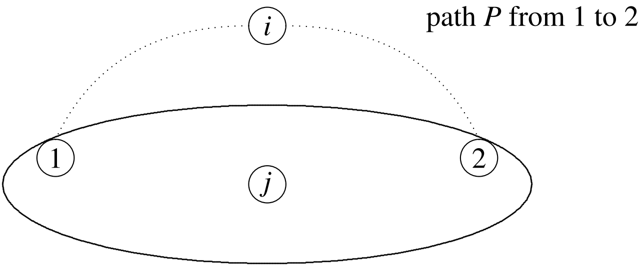
1. $\underset{i \in P}{\text{Average}}\, V_{ij} \geq \frac{1}{2}(V_{1j} + V_{2j})$ for all j in $G \setminus P$.

Proof outline



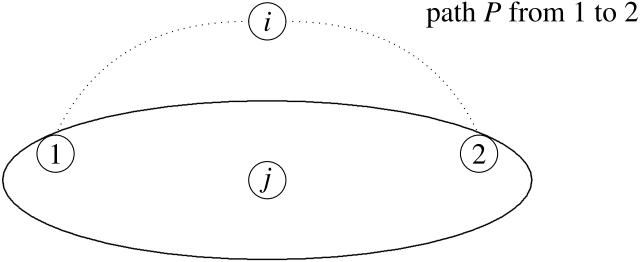
1. $\underset{i \in P}{\text{Average}}\, V_{ij} \geq \frac{1}{2}(V_{1j} + V_{2j})$ for all j in $G \setminus P$.
2. $\underset{\text{all } i,\, j}{\text{Average}}\, V_{ij} \geq \underset{i,\, j \in G_0}{\text{Average}}\, V_{ij}$

Proof outline



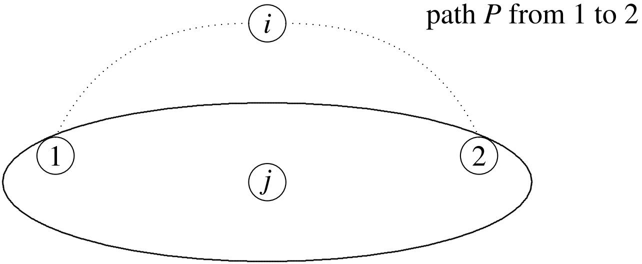
1. $\underset{i \in P}{\text{Average}}\, V_{ij} \geq \frac{1}{2}(V_{1j} + V_{2j})$ for all j in $G \setminus P$.
2. $\underset{\text{all } i,\, j}{\text{Average}}\, V_{ij} \geq \underset{i,\, j \in G_0}{\text{Average}}\, V_{ij}$
3. Inserting $n-1$ vertices into all edges of G_0 multiplies all original variances by n.

Proof outline



1. $\underset{i \in P}{\text{Average}}\, V_{ij} \geq \frac{1}{2}(V_{1j} + V_{2j})$ for all j in $G \setminus P$.
2. $\underset{\text{all } i,\, j}{\text{Average}}\, V_{ij} \geq \underset{i,\, j \in G_0}{\text{Average}}\, V_{ij}$
3. Inserting $n-1$ vertices into all edges of G_0 multiplies all original variances by n.
4. Large enough $v \implies$ large enough $n \implies \underset{\text{all } i,\, j}{\text{Average}}\, V_{ij} \geq 2$.

Proof outline



1. $\underset{i \in P}{\text{Average}}\, V_{ij} \geq \frac{1}{2}(V_{1j} + V_{2j})$ for all j in $G \setminus P$.
2. $\underset{\text{all } i,\, j}{\text{Average}}\, V_{ij} \geq \underset{i,\, j \in G_0}{\text{Average}}\, V_{ij}$
3. Inserting $n - 1$ vertices into all edges of G_0 multiplies all original variances by n.
4. Large enough $v \implies$ large enough $n \implies \underset{\text{all } i,\, j}{\text{Average}}\, V_{ij} \geq 2$.
5. In the star $K_{1,v-1}$, we have $\underset{\text{all } i,\, j}{\text{Average}}\, V_{ij} = \dfrac{2(v-1)}{v} \leq 2$.