Linear Algebraic methods in Combinatorics

Sivaramakrishnan S

Dept of Mathematics IIT Bombay

18 Apr 2023, BITS, Goa

Sivaramakrishnan S (IIT-B)

Linear Algebra in Combinatorics

18 Apr 2023, BITS Goa 1 / 26

A tale of four cities...

Sivaramakrishnan S (IIT-B)

< □ > < 同 >

æ

A tale of four cities...

Even City :

Sivaramakrishnan S (IIT-B)

▶ < ∃ >

< □ > < 同 >

æ

Even City : Has n = 40 inhabitants. They keep forming clubs.

э

Even City : Has n = 40 inhabitants. They keep forming clubs. How many clubs can there be if no two clubs have the same set of people?

Dictatorial rules :

Each club has an even number of people in it.

Dictatorial rules :

- Each club has an even number of people in it.
- 2 Every pair of clubs has an even number of people.

Dictatorial rules :

- Each club has an even number of people in it.
- 2 Every pair of clubs has an even number of people.
- In two clubs have the same set of members.

Even City

3

*ロト *個ト * ヨト * ヨト

Question : How many clubs can there now be?

э

Question : How many clubs can there now be? Easy to form $2^{20} > 10^6$ clubs. Quite big for a city with just 40 inhabitants!

Question : How many clubs can there now be? Easy to form $2^{20}>10^6$ clubs. Quite big for a city with just 40 inhabitants! Non-trivial fact : For any set of $\ell<2^{20}$ clubs, we can add one more club.

The second city - Odd City

Sivaramakrishnan S (IIT-B)

▶ < ∃ >

< □ > < 同 >

æ

Again has 40 inhabitants.

э

Each club has an odd number of people in it.

- Each club has an odd number of people in it.
- 2 Every pair of clubs has an odd number of people.

- Each club has an odd number of people in it.
- 2 Every pair of clubs has an odd number of people.
- So two clubs can have the same set of people.

- Each club has an odd number of people in it.
- 2 Every pair of clubs has an odd number of people.
- O No two clubs can have the same set of people.
- **Question :** How many clubs can there be?

Odd city

Sivaramakrishnan S (IIT-B)

æ

*ロト *個ト * ヨト * ヨト

Easy to form $2^{19}=524288\geq 1/2\cdot 10^6$ clubs.

э

▶ < ∃ >

Easy to form $2^{19} = 524288 \geq 1/2 \cdot 10^6$ clubs. Council members are certainly not happy.

э

The third city - Odd-Even city

< □ > < 同 >

æ

Again has 40 inhabitants.

э

Each club has an odd number of people in it.

- Each club has an odd number of people in it.
- 2 Every pair of clubs has an even number of people.

- Each club has an odd number of people in it.
- 2 Every pair of clubs has an even number of people.
- Question : How many clubs can there be?

Odd-Even city

Sivaramakrishnan S (IIT-B)

æ

<ロト < 四ト < 三ト < 三ト

Easy to see that there can be 40 clubs.

э

Easy to see that there can be 40 clubs. **Solution 1:** Each person is a club of one.

Easy to see that there can be 40 clubs. **Solution 1:** Each person is a club of one. **Solution 2:** Each club has 39 people, *i*-th club does not have the *i*-th person. Easy to see that there can be 40 clubs. **Solution 1:** Each person is a club of one. **Solution 2:** Each club has 39 people, *i*-th club does not have the *i*-th person.

Thus, there can be at least 40 clubs.

Easy to see that there can be 40 clubs. **Solution 1:** Each person is a club of one. **Solution 2:** Each club has 39 people, *i*-th club does not have the *i*-th person. Thus, there can be at least 40 clubs. **Question :** Can there be more?

Odd-Even city

Sivaramakrishnan S (IIT-B)

æ

<ロト < 四ト < 三ト < 三ト

There can be at most 40 clubs!

イロト イヨト イヨト イヨト

æ

There can be at most 40 clubs!

Proof:

イロト イヨト イヨト

æ

There can be at most 40 clubs!

Proof: Let there be *t* clubs, C_1, C_2, \ldots, C_t . Take the indicator vector $x_i \in \{0, 1\}^{40}$ of club C_i . The x_i 's lie in a 40 dimensional space.

э

There can be at most 40 clubs!

Proof : Let there be *t* clubs, C_1, C_2, \ldots, C_t . Take the indicator vector $x_i \in \{0, 1\}^{40}$ of club C_i . The x_i 's lie in a 40 dimensional space. They are linearly independent over \mathbb{Z}_2 . (This is a strong statement.)

There can be at most 40 clubs!

Proof: Let there be *t* clubs, C_1, C_2, \ldots, C_t . Take the indicator vector $x_i \in \{0, 1\}^{40}$ of club C_i . The x_i 's lie in a 40 dimensional space. They are linearly independent over \mathbb{Z}_2 . (This is a strong statement.) In a vector space of dimension 40, there cannot be more than 40 linearly independent vectors.

Proof of theorem

Sivaramakrishnan S (IIT-B)

イロト イヨト イヨト

æ

Note that $x_i \cdot x_j$ is odd (non zero in \mathbb{Z}_2) iff i = j.

э

Note that $x_i \cdot x_j$ is odd (non zero in \mathbb{Z}_2) iff i = j. If $\lambda_1 \cdot x_1 + \lambda_2 \cdot x_2 + \cdots + \lambda_t \cdot x_t = 0$

э

Note that $x_i \cdot x_j$ is odd (non zero in \mathbb{Z}_2) iff i = j. If $\lambda_1 \cdot x_1 + \lambda_2 \cdot x_2 + \cdots + \lambda_t \cdot x_t = 0$ Then take dot product with x_i to get $\lambda_i = 0$. Note that $x_i \cdot x_j$ is odd (non zero in \mathbb{Z}_2) iff i = j. If $\lambda_1 \cdot x_1 + \lambda_2 \cdot x_2 + \cdots + \lambda_t \cdot x_t = 0$ Then take dot product with x_i to get $\lambda_i = 0$. Do this for all *i*. Thus all λ_i 's are zero. Hence the x_i 's are linearly independent.

Another proof

Sivaramakrishnan S (IIT-B)

3

メロト メポト メヨト メヨト

Let there be t clubs, C_1, C_2, \ldots, C_t . Recall x_i is the indicator (column) vector of the club C_i .

Let there be t clubs, C_1, C_2, \ldots, C_t . Recall x_i is the indicator (column) vector of the club C_i . Take the 40 × t matrix M whose *i*-th column is x_i . Let there be t clubs, C_1, C_2, \ldots, C_t . Recall x_i is the indicator (column) vector of the club C_i . Take the 40 × t matrix M whose *i*-th column is x_i . Claim : rank_{Z₂}(M) = t. Let there be t clubs, C_1, C_2, \ldots, C_t . Recall x_i is the indicator (column) vector of the club C_i . Take the 40 × t matrix M whose *i*-th column is x_i . Claim : rank_{Z2}(M) = t. Proof : Let $N = M^T M$. Let there be t clubs, C_1, C_2, \ldots, C_t . Recall x_i is the indicator (column) vector of the club C_i .

Take the 40 \times t matrix M whose *i*-th column is x_i .

Claim : rank_{$$\mathbb{Z}_2$$}(*M*) = *t*.

Proof : Let
$$N = M^T M$$
.

Fact : For any field \mathbb{F} , rank_{\mathbb{F}}(AB) \leq min(rank_{\mathbb{F}}(A), rank_{\mathbb{F}}(B))

Another proof - cont'd

Sivaramakrishnan S (IIT-B)

▶ < ∃ >

< □ > < 同 >

æ

Thus, $\operatorname{rank}_{\mathbb{Z}_2}(N) \leq \operatorname{rank}_{\mathbb{Z}_2}(M)$.

3

Thus, $\operatorname{rank}_{\mathbb{Z}_2}(N) \leq \operatorname{rank}_{\mathbb{Z}_2}(M)$. Because of our odd-even condition, over \mathbb{Z}_2 , $N = I_{t \times t}$. Thus, $\operatorname{rank}_{\mathbb{Z}_2}(N) \leq \operatorname{rank}_{\mathbb{Z}_2}(M)$. Because of our odd-even condition, over \mathbb{Z}_2 , $N = I_{t \times t}$. Thus $t \leq \operatorname{rank}_{\mathbb{Z}_2}(M)$. Thus, $\operatorname{rank}_{\mathbb{Z}_2}(N) \leq \operatorname{rank}_{\mathbb{Z}_2}(M)$. Because of our odd-even condition, over \mathbb{Z}_2 , $N = I_{t \times t}$. Thus $t \leq \operatorname{rank}_{\mathbb{Z}_2}(M)$. We know that $\operatorname{rank}_{\mathbb{Z}_2}(M) \leq \min(t, 40)$. Thus, $\operatorname{rank}_{\mathbb{Z}_2}(N) \leq \operatorname{rank}_{\mathbb{Z}_2}(M)$. Because of our odd-even condition, over \mathbb{Z}_2 , $N = I_{t \times t}$. Thus $t \leq \operatorname{rank}_{\mathbb{Z}_2}(M)$. We know that $\operatorname{rank}_{\mathbb{Z}_2}(M) \leq \min(t, 40)$. i.e. $t \leq \operatorname{rank}_{\mathbb{Z}_2}(M) \leq \min(t, 40)$ and so $t \leq 40$.

Sivaramakrishnan S (IIT-B)

メロト メポト メモト メモト

æ

Recall: For even city, if there were $\ell < 2^{20}$ clubs, we could add another club without violating maximality.

Recall: For even city, if there were $\ell < 2^{20}$ clubs, we could add another club without violating maximality. **Question:** Is the same true for Odd-even town?

The fourth city...

Sivaramakrishnan S (IIT-B)

→ < ∃ →</p>

 $\langle \Box \rangle \langle \Box \rangle$

æ

Reverse City : Again has 40 inhabitants.

- Each club has an even number of people in it.
- 2 Every pair of clubs has an odd number of people.

Reverse City : Again has 40 inhabitants.

- Each club has an even number of people in it.
- 2 Every pair of clubs has an odd number of people.

Question : How many clubs can there be? Think it over...

Same intersection size

Sivaramakrishnan S (IIT-B)

< □ > < 同 >

æ

∃ ► < Ξ</p>

Theorem 2

If S_1, S_2, \ldots, S_t are distinct subsets of [n] such that for all $i \neq j$, $|S_i \cap S_j| = \ell$, then $t \leq n$.

Theorem 2

If S_1, S_2, \ldots, S_t are distinct subsets of [n] such that for all $i \neq j$, $|S_i \cap S_j| = \ell$, then $t \leq n$.

Proof: If one S_i has size ℓ , then all other sets contain S_i and are mutually disjoint outside S_i . Thus $t \le n - \ell \le n$.

Theorem 2

If S_1, S_2, \ldots, S_t are distinct subsets of [n] such that for all $i \neq j$, $|S_i \cap S_j| = \ell$, then $t \leq n$.

Proof: If one S_i has size ℓ , then all other sets contain S_i and are mutually disjoint outside S_i . Thus $t \le n - \ell \le n$. If no set S_i has size ℓ , then $|S_i| > \ell$ for all i.

Theorem 2

If S_1, S_2, \ldots, S_t are distinct subsets of [n] such that for all $i \neq j$, $|S_i \cap S_j| = \ell$, then $t \leq n$.

Proof: If one S_i has size ℓ , then all other sets contain S_i and are mutually disjoint outside S_i . Thus $t \le n - \ell \le n$. If no set S_i has size ℓ , then $|S_i| > \ell$ for all i. Let M be the $n \times t$ matrix and let $N = M^T M$.

Proof - cont'd

Sivaramakrishnan S (IIT-B)

3

 $\operatorname{rank}_{\mathbb{Q}}(N) = t.$

Sivaramakrishnan S (IIT-B)

Linear Algebra in Combinatorics

18 Apr 2023, BITS Goa 15 / 26

æ

<ロト < 四ト < 三ト < 三ト

 $\operatorname{rank}_{\mathbb{Q}}(N) = t.$

Stronger claim: The matrix *N* is positive definite.

3

ヨト イヨト

< A

 $\operatorname{rank}_{\mathbb{Q}}(N) = t.$

Stronger claim: The matrix *N* is positive definite. Let $K = \text{diag}[|S_i| - \ell]$.

э

▶ < ∃ >

 $\operatorname{rank}_{\mathbb{Q}}(N) = t.$

Stronger claim: The matrix *N* is positive definite. Let $K = \text{diag}[|S_i| - \ell]$. It is easy to see that $N = \ell J + K$.

Ramsey Theory

Sivaramakrishnan S (IIT-B)

3

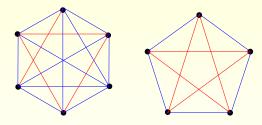
*ロト *個ト * ヨト * ヨト

Theorem 3

In ANY colouring of the edges of the complete graph on 6 vertices with two colours red and blue, there is a monochromatic triangle. The same is NOT true if we were to colour the complete graph on 5 vertices.

Theorem 3

In ANY colouring of the edges of the complete graph on 6 vertices with two colours red and blue, there is a monochromatic triangle. The same is NOT true if we were to colour the complete graph on 5 vertices.



Cliques in graphs

Sivaramakrishnan S (IIT-B)

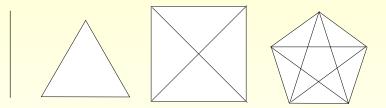
æ

(ロ) (日) (日) (日) (日)

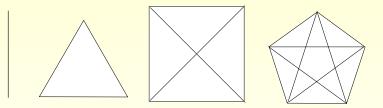
A clique (or complete graph) on n vertices is a graph where each pair of edges is present.

э

A clique (or complete graph) on n vertices is a graph where each pair of edges is present.



A clique (or complete graph) on n vertices is a graph where each pair of edges is present.



Call R(rd, bl) as the minimum number of vertices such that ANY colouring of the edges of the complete graph on R(rd, bl) vertices has a RED clique of size rd or a BLUE clique of size bl.

Ramsey Theory

Sivaramakrishnan S (IIT-B)

3

*ロト *個ト * ヨト * ヨト

R(3,3) = 6.

3

<ロト < 回 > < 回 > < 回 > < 回 >

$$R(3,3) = 6. R(2,t) = ?.$$

3

$$R(3,3) = 6. R(2,t) = ?.$$

Question 2

Does R(p,q) exist for all p and q?

크

ヨト イヨト

Image: A matrix

$$R(3,3) = 6. R(2,t) = ?.$$

Question 2

Does R(p,q) exist for all p and q?

Theorem 4 (Ramsey, 1927)

For all p, q the number R(p, q) is finite.

3

★ Ξ ► < Ξ ►</p>

Image: A matrix

$$R(3,3) = 6. R(2,t) = ?.$$

Question 2

Does R(p,q) exist for all p and q?

Theorem 4 (Ramsey, 1927)

For all p, q the number R(p, q) is finite.

We don't know exact values of R(p, q) for arbitrary p and q.

Sivaramakrishnan S (IIT-B)

3

▲口 ▶ ▲圖 ▶ ▲ 臣 ▶ ▲ 臣 ▶

We have seen R(3,3) = 6.

æ

(ロ) (日) (日) (日) (日)

We have seen R(3,3) = 6. It is known that R(4,4) = 18.

ヨト イヨト

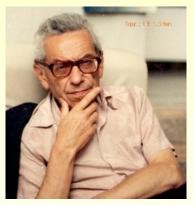
< □ > < 同 >

э

We have seen R(3,3) = 6. It is known that R(4,4) = 18. Paul Erdös joked about these numbers

э

We have seen R(3,3) = 6. It is known that R(4,4) = 18. Paul Erdös joked about these numbers



Paul Erdös (1913 – 1996).

We have seen R(3,3) = 6. It is known that R(4,4) = 18. Paul Erdös joked about these numbers



Paul Erdös (1913 – 1996).

Can we get bounds on these Ramsey Numbers?

Bounds on Ramsey numbers

Sivaramakrishnan S (IIT-B)

э

Theorem 5 (Erdös, 1960's)

 $(\sqrt{2})^k \leq R(k,k) \leq 4^k$.

Sivaramakrishnan S (IIT-B)

< A

э

Theorem 5 (Erdös, 1960's)

 $(\sqrt{2})^k \leq R(k,k) \leq 4^k$. Alas, the proofs of Erdös are probabilistic. We do not know an explicit family of graphs on $(\sqrt{2})^k$ vertices and a proof that the family is a Ramsey graph.

An explicit lower bound

Sivaramakrishnan S (IIT-B)

▶ < ∃ >

< □ > < 同 >

æ

 $R(k,k) \geq (k-1)^2.$

크

 $R(k,k) \geq (k-1)^2.$

Theorem 7 (Nagy, 1972)

 $R(k,k)\geq \Omega(k^3)$

Sivaramakrishnan S (IIT-B)

3

∃ ► < ∃ ►</p>

 $R(k,k) \geq (k-1)^2.$

Theorem 7 (Nagy, 1972)

 $R(k,k) \geq \Omega(k^3)$

Consider a universe U with k elements.

3

ヨト イヨト

 $R(k,k) \geq (k-1)^2.$

Theorem 7 (Nagy, 1972)

 $R(k,k) \geq \Omega(k^3)$

Consider a universe U with k elements. Nagy's graph has all possible 3-subsets of U as vertices.

ヨト イヨト

 $R(k,k) \geq (k-1)^2.$

Theorem 7 (Nagy, 1972)

 $R(k,k) \geq \Omega(k^3)$

Consider a universe U with k elements. Nagy's graph has all possible 3-subsets of U as vertices. The edge connecting R and S is coloured blue iff $|R \cap S| = 1$.

 $R(k,k) \geq (k-1)^2.$

Theorem 7 (Nagy, 1972)

 $R(k,k) \ge \Omega(k^3)$

Consider a universe U with k elements. Nagy's graph has all possible 3-subsets of U as vertices. The edge connecting R and S is coloured blue iff $|R \cap S| = 1$. **Claim :** This graph has no monochromatic clique of size k.

A B F A B F

Sivaramakrishnan S (IIT-B)

3

<ロト < 四ト < 三ト < 三ト

Argument against blue cliques

ヨト イヨト

< □ > < 同 >

æ

Argument against blue cliques Blue clique implies sets S_1, S_2, \dots, S_t such that for all $i \neq j$, $|S_i \cap S_j| = 1$.

э

ヨト イヨト

Image: A matrix

Argument against blue cliques Blue clique implies sets S_1, S_2, \dots, S_t such that for all $i \neq j$, $|S_i \cap S_j| = 1$. Use the same intersection size theorem!

Argument against blue cliques Blue clique implies sets S_1, S_2, \dots, S_t such that for all $i \neq j$, $|S_i \cap S_j| = 1$. Use the same intersection size theorem! Argument against red cliques

Argument against blue cliques Blue clique implies sets S_1, S_2, \dots, S_t such that for all $i \neq j$, $|S_i \cap S_j| = 1$. Use the same intersection size theorem! Argument against red cliques Red clique implies sets S_1, S_2, \dots, S_t such that for all $i \neq j$, $|S_i \cap S_j| = 0, 2$, but $|S_i| = 3$. Argument against blue cliques Blue clique implies sets S_1, S_2, \dots, S_t such that for all $i \neq j$, $|S_i \cap S_j| = 1$. Use the same intersection size theorem! Argument against red cliques Red clique implies sets S_1, S_2, \dots, S_t such that for all $i \neq j$, $|S_i \cap S_j| = 0, 2$, but $|S_i| = 3$. Use Odd-even town theorem! Argument against blue cliques Blue clique implies sets S_1, S_2, \dots, S_t such that for all $i \neq j$, $|S_i \cap S_j| = 1$. Use the same intersection size theorem! Argument against red cliques Red clique implies sets S_1, S_2, \dots, S_t such that for all $i \neq j$, $|S_i \cap S_j| = 0, 2$, but $|S_i| = 3$. Use Odd-even town theorem!

Thus there is no monochromatic clique of size k!!

Let K_n be the complete graph on *n* vertices. We want to cover all the edges of K_n using complete bipartite graphs K_{S_i,T_i} such that each edge of K_n occurs in precisely one K_{S_i,T_i} and use the minimum number of complete bipartite graphs.

Let K_n be the complete graph on *n* vertices. We want to cover all the edges of K_n using complete bipartite graphs K_{S_i,T_i} such that each edge of K_n occurs in precisely one K_{S_i,T_i} and use the minimum number of complete bipartite graphs.

Theorem 8 (Graham-Pollak)

The minimum number of complete bipartite graphs needed to cover the edges of K_n (as a disjoint union) is n - 1.

Sivaramakrishnan S (IIT-B)

æ

▲□▶ ▲圖▶ ▲臣▶ ▲臣▶

Proof: Assume that the vertex set of K_n is $\{1, 2, ..., n\}$. Associate a polynomial $P_G(x_1, ..., x_n) = \sum_{e \in E(G), e = \{i,j\}} x_i x_j$ to the graph G.

Proof: Assume that the vertex set of K_n is $\{1, 2, ..., n\}$. Associate a polynomial $P_G(x_1, ..., x_n) = \sum_{e \in E(G), e = \{i, j\}} x_i x_j$ to the graph G. Suppose K_{S_i, T_i} for $1 \le i \le q$ covers the edges of K_n , and let $q \le n - 2$.

Sivaramakrishnan S (IIT-B)

Proof: Assume that the vertex set of K_n is $\{1, 2, ..., n\}$. Associate a polynomial $P_G(x_1, ..., x_n) = \sum_{e \in E(G), e = \{i,j\}} x_i x_j$ to the graph G. Suppose K_{S_i, T_i} for $1 \le i \le q$ covers the edges of K_n , and let $q \le n - 2$. We clearly have $P_{K_n}(x_1, x_2, ..., x_n) = \sum_{i=1}^q P_{K_{S_i, T_i}}(x_1, x_2, ..., x_n)$.

It is easy to see that

$$P_{K_n}(x_1, x_2, \dots, x_n) = \sum_{1 \le i < j \le n} x_i x_j = \frac{1}{2} \left[(\sum_{i=1}^n x_i)^2 - \sum_{i=1}^n x_i^2 \right].$$

▲□▶ ▲圖▶ ▲国▶ ▲国▶ 二国

It is easy to see that

$$P_{K_n}(x_1, x_2, \dots, x_n) = \sum_{1 \le i < j \le n} x_i x_j = \frac{1}{2} \left[(\sum_{i=1}^n x_i)^2 - \sum_{i=1}^n x_i^2 \right].$$

and that $P_{K_{S_a,T_a}}(x_1, x_2, \dots, x_n) = (\sum_{i \in S_a} x_i) (\sum_{j \in T_a} x_j).$

◆□ > ◆圖 > ◆臣 > ◆臣 > □ 臣

It is easy to see that

$$P_{K_n}(x_1, x_2, ..., x_n) = \sum_{1 \le i < j \le n} x_i x_j = \frac{1}{2} \left[(\sum_{i=1}^n x_i)^2 - \sum_{i=1}^n x_i^2 \right].$$
and that $P_{K_{S_a,T_a}}(x_1, x_2, ..., x_n) = (\sum_{i \in S_a} x_i) (\sum_{j \in T_a} x_j).$
Consider the linear homogeneous system of equations $\sum_{i \in S_k} x_i = 0$ for
 $1 \le k \le q$ and $\sum_{i=1}^n x_i = 0.$

イロト イ部ト イヨト イヨト 二日

It is easy to see that $P_{K_n}(x_1, x_2, \dots, x_n) = \sum_{1 \le i < j \le n} x_i x_j = \frac{1}{2} \left[\left(\sum_{i=1}^n x_i \right)^2 - \sum_{i=1}^n x_i^2 \right].$ and that $P_{K_{S_a, T_a}}(x_1, x_2, \dots, x_n) = \left(\sum_{i \in S_a} x_i \right) \left(\sum_{j \in T_a} x_j \right).$ Consider the linear homogeneous system of equations $\sum_{i \in S_k} x_i = 0$ for $1 \le k \le q$ and $\sum_{i=1}^n x_i = 0.$ Since $q \le n-2$, this system has *n* variables and at most n-1 equations. Thus, over \mathbb{R} , there is a non-zero solution $(z_1, z_2, \dots, z_n).$ It is easy to see that $P_{K_n}(x_1, x_2, \dots, x_n) = \sum_{1 \le i < j \le n} x_i x_j = \frac{1}{2} \left[(\sum_{i=1}^n x_i)^2 - \sum_{i=1}^n x_i^2 \right].$ and that $P_{K_{S_a,T_a}}(x_1, x_2, \dots, x_n) = (\sum_{i \in S_a} x_i) (\sum_{j \in T_a} x_j).$ Consider the linear homogeneous system of equations $\sum_{i \in S_k} x_i = 0$ for $1 \le k \le q$ and $\sum_{i=1}^n x_i = 0.$ Since $q \le n-2$, this system has *n* variables and at most n-1 equations. Thus, over \mathbb{R} , there is a non-zero solution $(z_1, z_2, \dots, z_n).$ This solution violates $P_{K_n}(x_1, x_2, \dots, x_n) = \sum_{i=1}^q P_{K_{S_i,T_i}}(x_1, x_2, \dots, x_n).$ This gives us a contradiction. Questions/ Comments???

▶ < ∃ >

< □ > < 同 >

æ