## **Generalized Cauchy-Schwarz Inequality**

Vincent Tam

Notes for <u>a journal article on this inequality</u>

For any positive real numbers

$$\underbrace{\left(\prod_{k=1}^{n}a_{i}\right)^{1/n}}_{\text{geometric mean}} \leq \underbrace{\frac{1}{n}\left(\sum_{k=1}^{n}a_{i}\right)}_{\text{arithmetic mean}}$$

of  $a_1, ..., a_n$ 

Equality holds if and only if all

$$a_1 = \dots = a_n$$
.

of  $a_1, ..., a_n$ 

For any positive real numbers  $a_1, ..., a_n$ ,

$$\underbrace{\left(\prod_{k=1}^{n}a_i\right)^{1/n}}_{\text{geometric mean of }a_1,\dots,a_n} \leq \underbrace{\frac{1}{n}\left(\sum_{k=1}^{n}a_i\right)}_{\text{arithmetic mean of }a_1,\dots,a_n}$$

Equality holds if and only if all  $a_1 = \cdots = a_n$ .

*Proof.* (by replacement) If all  $a_i$ 's are equal, then it's trivial.

For any positive real numbers  $a_1, ..., a_n$ ,

$$\underbrace{\left(\prod_{k=1}^{n}a_i\right)^{1/n}}_{\text{geometric mean of }a_1,\dots,a_n} \leq \underbrace{\frac{1}{n}\left(\sum_{k=1}^{n}a_i\right)}_{\text{arithmetic mean of }a_1,\dots,a_n}$$

Equality holds if and only if all  $a_1 = \cdots = a_n$ .

*Proof.* (by replacement) If all  $a_i$ 's are equal, then it's trivial. If not, let

$$\alpha = \underbrace{\mathrm{AM}(a_1, ..., a_n)}_{\text{arithmetic mean of } a_1, ..., a_n}.$$

For any positive real numbers  $a_1, ..., a_n$ ,

$$\underbrace{\left(\prod_{k=1}^{n}a_i\right)^{1/n}}_{\text{geometric mean}} \leq \underbrace{\frac{1}{n}\left(\sum_{k=1}^{n}a_i\right)}_{\text{arithmetic mean}}$$
 of  $a_1,\dots,a_n$ 

Equality holds if and only if all  $a_1 = \cdots = a_n$ .

*Proof.* (by replacement) If all  $a_i$ 's are equal, then it's trivial. If not, let

$$\alpha = \underbrace{\mathrm{AM}(a_1, ..., a_n)}_{\text{arithmetic mean of } a_1, ..., a_n}.$$

Then there exists  $\underbrace{i \text{ and } j}_{\text{indices}}$  such that

$$a_i < \alpha < a_j$$

For any positive real numbers

$$\underbrace{\left(\prod_{k=1}^{n}a_{i}\right)^{1/n}}_{\text{geometric mean of }a_{1},...,a_{n}}\leq\underbrace{\frac{1}{n}\left(\sum_{k=1}^{n}a_{i}\right)}_{\text{arithmetic mean of }a_{1},...,a_{n}}$$

Equality holds if and only if all  $a_1 = \cdots = a_n$ .

$$\alpha = \underbrace{\mathrm{AM}(a_1,...,a_n)}_{\text{arithmetic mean of } a_1,...,a_n}.$$
 Then there exists  $\underbrace{i \text{ and } j}_{\text{indices}}$  such that

*Proof.* (by replacement) If not, let

 $a_i < \alpha < a_j$ . (' $\alpha$ ' comes from "arithmetic mean")

For any positive real numbers

$$\underbrace{\left(\prod_{k=1}^{n}a_{i}\right)^{1/n}}_{\text{geometric mean of }a_{1},...,a_{n}}\leq\underbrace{\frac{1}{n}\left(\sum_{k=1}^{n}a_{i}\right)}_{\text{arithmetic mean of }a_{1},...,a_{n}}$$

Equality holds if and only if all  $a_1 = \cdots = a_n$ .

*Proof.* (by replacement) If not, let

$$\alpha = \underbrace{\mathrm{AM}(a_1,...,a_n)}_{\text{arithmetic mean of } a_1,...,a_n}.$$

Then there exists  $\underbrace{i \text{ and } j}_{\text{indices}}$  such that

$$\underbrace{a_{i} < \alpha < a_{j}}_{a_{i} \leftarrow \alpha}, a_{j} \leftarrow a_{i} + a_{j} - \alpha,$$

$$\underbrace{a_{i} \leftarrow \alpha}_{a_{i} \text{ replaced by } \alpha}$$

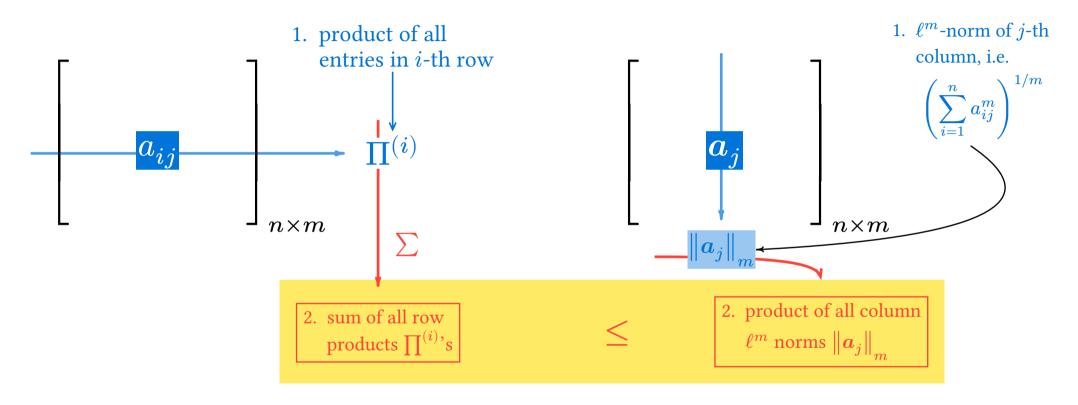
so after this replacement process,

old AM = new AM, but old GM < new GM because 
$$a_i a_j < \alpha (a_i + a_j - \alpha)$$
  $\Leftrightarrow (\alpha - a_i)(\alpha - a_j) < 0.$  After this replacement, in the new AM we have at least one less number not equal to  $\alpha$ , This process can be repeated until all  $a_i$ 's are equal (to  $\alpha$ ), then final GM =  $\alpha$ .

#### Hence

initial 
$$AM = \alpha = \text{final GM}$$
  
  $\geq \text{new GM} > \text{initial GM}.$ 

#### **Generalized Cauchy-Schwarz Inequality**



# Question: How to remember this inequality?

### Question: How to remember this inequality?

- ' $\longrightarrow$ ' linked with ' $\prod$ ' as 'H' in "Horizontal" looks like ' $\prod$ '
- ' $\downarrow$ ' linked with ' $\sum$ ' as two 'V's ("Vertical") looks like ' $\lessgtr$ ' Orders of arrows in each side:
- LHS: natural reading order ("less strange") RHS: "more strange" reading order
  - 1. "left  $\longrightarrow$  right" first
  - top then bottom

- - top first bottom
  - 2. then "left  $\longrightarrow$  right"

Overall strategy is similar to the proof of Hölder's inequality. Observe that in the target inequality

$$\sum_{i=1}^{n} \prod_{j=1}^{m} a_{ij} \leq \prod_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ij}^{m}\right)^{1/m},$$

if we multiply (the entries of) the j-th column by a positive constant k (i.e. for each  $i \in \{1, ..., n\}$  and a particular fixed  $j \in \{1, ..., n\}$ ,  $a_{ij} \leftarrow ka_{ij}$ ), each side of the above inequality is also multiplied by k.

Observe that in the target inequality

$$\sum_{i=1}^{n} \prod_{j=1}^{m} a_{ij} \le \prod_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ij}^{m}\right)^{1/m},$$

WLOG (without loss of generality), we can assume that the j-th column is normalized, i.e.  $\|a_j\|_m=1$ .

Observe that in the target inequality

$$\sum_{i=1}^{n} \prod_{j=1}^{m} a_{ij} \le \prod_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ij}^{m}\right)^{1/m},$$

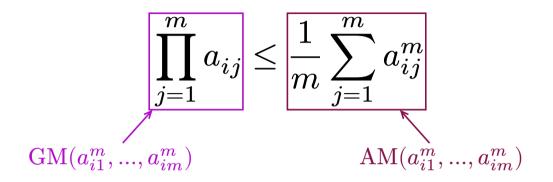
WLOG (without loss of generality), we can assume that the j-th column is normalized, i.e.  $\|a_j\|_m=1$ . The same goes for the remaining columns.

Observe that in the target inequality

$$\sum_{i=1}^{n} \prod_{j=1}^{m} a_{ij} \le \prod_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ij}^{m}\right)^{1/m},$$

WLOG (without loss of generality), we can assume that the j-th column is normalized, i.e.  $\|a_j\|_m = 1$ . The same goes for the remaining columns. Then the RHS becomes 1.

### Proof step 2: AM-GM on each row product



#### Guide:

- 1. ' $\sum_{i} \prod_{j} a_{ij}$ ' in LHS of <u>previous slide</u> seems hard, so tackle each row product  $\prod_{j} a_{ij}$  with AM-GM first. A "product" reminds us of "geometric mean".
- 2. The power 'm' (in superscript) and the ' $\sum$ ' in RHS of <u>previous slide</u> seems to be an arithmetic mean.

In target LHS, we have ' $\sum_i$ ' on the left of ' $\prod_j$ ', so take ' $\sum_i$ ' on both sides of the inequality in the <u>previous step</u>.

$$\sum_{i=1}^{n} \prod_{j=1}^{m} a_{ij} \le \sum_{i=1}^{n} \frac{1}{m} \sum_{j=1}^{m} a_{ij}^{m}$$

In target LHS, we have  $\sum_{i}$  on the left of  $\prod_{j}$ , so take  $\sum_{i}$  on both sides of the inequality in the <u>previous step</u>.

$$\sum_{i=1}^{n} \prod_{j=1}^{m} a_{ij} \le \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij}^{m}$$

In target LHS, we have ' $\sum_i$ ' on the left of ' $\prod_j$ ', so take ' $\sum_i$ ' on both sides of the inequality in the <u>previous step</u>.

$$\sum_{i=1}^{n} \prod_{j=1}^{m} a_{ij} \le \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ij}^{m}$$

$$= \frac{1}{m} \sum_{j=1}^{m} \frac{\|\mathbf{a}_{j}\|_{m}^{m}}{\uparrow}$$

each column is normalized

In target LHS, we have ' $\sum_i$ ' on the left of ' $\prod_j$ ', so take ' $\sum_i$ ' on both sides of the inequality in the <u>previous step</u>.

$$\sum_{i=1}^{n} \prod_{j=1}^{m} a_{ij} \le \frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{m} a_{ij}^{m}$$

$$= \frac{1}{m} \sum_{j=1}^{m} \|\mathbf{a}_{j}\|_{m}^{m}$$

$$= 1$$

We've applied the AM–GM inequality to each row product, so equality holds if and only if all entries in each row are equal,

i.e.  $a_{i1}=\cdots=a_{im}$  for all (row index)  $i\in\{1,...,n\}$ .

We've applied the AM–GM inequality to each row product, so equality holds if and only if all entries in each row are equal,

i.e.  $a_{i1}=\cdots=a_{im}$  for all (row index)  $i\in\{1,...,n\}$ .

In this case, each corresponding component in any two distinct column vectors is equal.

i.e. for all (row index)  $i \in \{1, ..., n\}$  and (column indices)  $j, j' \in \{1, ..., m\}, a_{ij} = a_{ij'}$ .

We've applied the AM–GM inequality to each row product, so equality holds if and only if all entries in each row are equal,

i.e.  $a_{i1}=\cdots=a_{im}$  for all (row index)  $i\in\{1,...,n\}$ .

In this case, each corresponding component in any two distinct column vectors is equal.

i.e. for all (row index)  $i \in \{1, ..., n\}$  and (column indices)  $j, j' \in \{1, ..., m\}, a_{ij} = a_{ij'}$ . Two vectors are equal to each other if and only if each of their corresponding components are equal.

We've applied the AM–GM inequality to each row product, so equality holds if and only if all entries in each row are equal,

i.e.  $a_{i1}=\cdots=a_{im}$  for all (row index)  $i\in\{1,...,n\}$ .

In this case, each corresponding component in any two distinct column vectors is equal.

i.e. for all (row index)  $i \in \{1, ..., n\}$  and (column indices)  $j, j' \in \{1, ..., m\}, a_{ij} = a_{ij'}$ . Two vectors are equal to each other if and only if each of their corresponding components are equal.

Since we have normalized each column in the first step of our proof, we have equality holds if and only if all column vectors in the matrix are parallel to each other.

## Corollary: Carlson's inequality

$$\frac{\sum_{i=1}^{n} \sqrt[m]{\prod_{j=1}^{m} a_{ij}}}{n} \leq \sqrt[m]{\prod_{j=1}^{m} \frac{\sum_{i=1}^{n} a_{ij}}{n}}$$

## Corollary: Carlson's inequality

$$\frac{\sum_{i=1}^{n} \sqrt[m]{\prod_{j=1}^{m} a_{ij}}}{n} \leq \sqrt[m]{\prod_{j=1}^{m} \frac{\sum_{i=1}^{n} a_{ij}}{n}}$$

*Proof.* Apply the generalized Cauchy–Schwarz inequality to the matrix

$$\begin{bmatrix} a_{11}^{1/m}/n & \cdots & a_{1m}^{1/m}/n \\ \vdots & \ddots & \vdots \\ a_{n1}^{1/m}/n & \cdots & a_{nm}^{1/m}/n \end{bmatrix}.$$

9 / 17

## Application: avoid fractional powers

The figure in the slide for the generalized Cauchy–Schwarz inequality is often too difficult to apply on questions. In practice, we often take the m-th power on both sides to avoid fractional powers.

i.e. 
$$\left(\sum_{i=1}^{n} \prod_{j=1}^{m} a_{ij}\right)^{m} \leq \prod_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ij}^{m}\right)$$
.

The following question is a good example to illustrate how arranging terms in the form of a matrix can help organizing thoughts.

*Example.* Let  $0 < \theta < \pi/2$ . Find the minimum value of  $\frac{2}{\sin \theta} + \frac{3}{\cos \theta}$ .

Example. Let  $0 < \theta < \pi/2$ . Find the minimum value of  $\frac{2}{\sin \theta} + \frac{3}{\cos \theta}$ .

#### Discussion:

- 1. Constraint: Pythagorean identity  $\sin^2 \theta + \cos^2 \theta = 1$ .
- 2. Objective function should stay on RHS.
- 3. Tricky part:

$$egin{bmatrix} rac{2}{\sin heta} & \sin^2 heta \ rac{3}{\cos heta} & \cos^2 heta \end{bmatrix}.$$

*Example.* Let  $0 < \theta < \pi/2$ . Find the minimum value of  $\frac{2}{\sin \theta} + \frac{3}{\cos \theta}$ .

#### Discussion:

- 1. Constraint: Pythagorean identity  $\sin^2 \theta + \cos^2 \theta = 1$ .
- 2. Objective function should stay on RHS.
- 3. Tricky part:

$$\begin{bmatrix} rac{2}{\sin heta} & rac{2}{\sin heta} & \sin^2 heta \ rac{3}{\cos heta} & rac{3}{\cos heta} & \cos^2 heta \end{bmatrix}.$$

Example. Let  $0 < \theta < \pi/2$ . Find the minimum value of  $\frac{2}{\sin \theta} + \frac{3}{\cos \theta}$ .

#### Discussion:

- 1. Constraint: Pythagorean identity  $\sin^2 \theta + \cos^2 \theta = 1$ . doesn't match the (equality)
- 2. Objective function should stay on RHS.
- 3. Tricky part:

$$\begin{bmatrix} \frac{2}{\sin \theta} & \frac{2}{\sin \theta} & \sin^2 \theta \\ \frac{3}{\cos \theta} & \frac{3}{\cos \theta} & \cos^2 \theta \end{bmatrix}.$$

Problem: #col = m = 3, so

rightmost column norm (on RHS)

constraint in point 1.

*Example.* Let  $0 < \theta < \pi/2$ . Find the minimum value of  $\frac{2}{\sin \theta} + \frac{3}{\cos \theta}$ .

#### Discussion:

- 1. Constraint: Pythagorean identity  $\sin^2 \theta + \cos^2 \theta = 1$ .
- 2. Objective function should stay on RHS.
- 3. Tricky part:

$$\begin{bmatrix} \frac{2}{\sin \theta} & \frac{2}{\sin \theta} & \sin^2 \theta \\ \frac{3}{\cos \theta} & \frac{3}{\cos \theta} & \cos^2 \theta \end{bmatrix}.$$

Problem: #col = m = 3, so rightmost column norm (on RHS) doesn't match the (equality) constraint in point 1. Quickfix: adjust the power of each term by taking the m-th root

i.e. 
$$\begin{bmatrix} \left(\frac{2}{\sin\theta}\right)^{1/3} & \left(\frac{2}{\sin\theta}\right)^{1/3} & \sin^{2/3}\theta \\ \left(\frac{3}{\cos\theta}\right)^{1/3} & \left(\frac{3}{\cos\theta}\right)^{1/3} & \cos^{2/3}\theta \end{bmatrix}.$$

## Problem solving flow for minimization problems

It would be hard to get the <u>final matrix</u> at the first sight, so I suggest the following steps (to "get the row product right" first).

0. Get rid of fractional powers (by writing a logically equivalent inequality with only integer powers).

## Problem solving flow for minimization problems

It would be hard to get the <u>final matrix</u> at the first sight, so I suggest the following steps (to "get the row product right" first).

- 0. Get rid of fractional powers (by writing a logically equivalent inequality with only integer powers).
- 1. Identify the relevant equality constraint (, which is independent of the choice of matrix).

## Problem solving flow for minimization problems

It would be hard to get the <u>final matrix</u> at the first sight, so I suggest the following steps (to "get the row product right" first).

- 0. Get rid of fractional powers (by writing a logically equivalent inequality with only integer powers).
- 1. Identify the relevant equality constraint (, which is independent of the choice of matrix).
- 2. Identify each term in the objective function, and place each of them in a row. Now you have two columns.

- 0. Get rid of fractional powers (by writing a logically equivalent inequality with only integer powers).
- 1. Identify the relevant equality constraint (, which is independent of the choice of matrix).
- 2. Identify each term in the objective function, and place each of them in a row. Now you have two columns.
- 3. Make suitable amount of copies of columns, so that each row product are "balanced", i.e. your row products are constants/terms on the RHS of the target inequality.

- 0. Get rid of fractional powers (by writing a logically equivalent inequality with only integer powers).
- 1. Identify the relevant equality constraint (, which is independent of the choice of matrix).
- 2. Identify each term in the objective function, and place each of them in a row. Now you have two columns.
- 3. Make suitable amount of copies of columns, so that each row product are "balanced", i.e. your row products are constants/terms on the RHS of the target inequality.
- 4. Count the number of columns, and take this number as m.

- 0. Get rid of fractional powers (by writing a logically equivalent inequality with only integer powers).
- 1. Identify the relevant equality constraint (, which is independent of the choice of matrix).
- 2. Identify each term in the objective function, and place each of them in a row. Now you have two columns.
- 3. Make suitable amount of copies of columns, so that each row product are "balanced", i.e. your row products are constants/terms on the RHS of the target inequality.
- 4. Count the number of columns, and take this number as m.
- 5. Take the m-th root of each term in the matrix (, so that the rightmost column norm matches the equality constraint.)

- 0. Get rid of fractional powers (by writing a logically equivalent inequality with only integer powers).
- 1. Identify the relevant equality constraint (, which is independent of the choice of matrix).
- 2. Identify each term in the objective function, and place each of them in a row. Now you have two columns.
- 3. Make suitable amount of copies of columns, so that each row product are "balanced", i.e. your row products are constants/terms on the RHS of the target inequality.
- 4. Count the number of columns, and take this number as m.
- 5. Take the m-th root of each term in the matrix (, so that the rightmost column norm matches the equality constraint.)
- 6. Apply the generalized Cauchy–Schwarz inequality to the matrix.

- 0. Get rid of fractional powers (by writing a logically equivalent inequality with only integer powers).
- 1. Identify the relevant equality constraint (, which is independent of the choice of matrix).
- 2. Identify each term in the objective function, and place each of them in a row. Now you have two columns.
- 3. Make suitable amount of copies of columns, so that each row product are "balanced", i.e. your row products are constants/terms on the RHS of the target inequality.
- 4. Count the number of columns, and take this number as m.
- 5. Take the m-th root of each term in the matrix (, so that the rightmost column norm matches the equality constraint.)
- 6. Apply the generalized Cauchy–Schwarz inequality to the matrix.
- 7. State the equality case.

#### Practice: generalization of <u>previous example</u>

*Exercise.* If 
$$a, b > 0, n \in \mathbb{N}, 0 < \theta < \pi/2$$
, show that 
$$\left(a^{2/(n+2)} + b^{2/(n+2)}\right)^{(n+2)2} \leq \frac{a}{\sin^n \theta} + \frac{b}{\cos^n \theta}.$$

# Practice: generalization of previous example

Exercise. If a, b > 0,  $n \in \mathbb{N}$ ,  $0 < \theta < \pi/2$ , show that

$$(a^{2/(n+2)} + b^{2/(n+2)})^{(n+2)2} \le \frac{a}{\sin^n \theta} + \frac{b}{\cos^n \theta}.$$

#### *Hint*:

- two columns of  $\begin{bmatrix} a/\sin^n \theta \\ b/\cos^n \theta \end{bmatrix}$  n columns of  $\begin{bmatrix} \cos^2 \theta \\ \sin^2 \theta \end{bmatrix}$

# Practice: Power Mean Inequality for integer power

*Exercise*. For any positive real numbers  $a_1, ..., a_n$  and positive integer p > 0, show that

$$\left(\frac{\sum_{i=1}^n a_i^p}{n}\right)^{1/p} \ge \frac{\sum_{i=1}^n a_i}{n}.$$

# Practice: Power Mean Inequality for integer power

Exercise. For any positive real numbers  $a_1, ..., a_n$  and positive integer p > 0, show that

$$\left(\frac{\sum_{i=1}^n a_i^p}{n}\right)^{1/p} \ge \frac{\sum_{i=1}^n a_i}{n}.$$

Solution. Apply the generalized Cauchy-Schwarz inequality to the matrix

$$\begin{bmatrix} a_1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_n & 1 & \cdots & 1 \end{bmatrix}$$

with p-1 columns of 1's.

#### Variation: Generalized <u>Titu's Lemma</u>

For any real numbers  $a_1, ..., a_n$ , positive real numbers  $b_1, ..., b_n$ , positive integers  $m, k \in \mathbb{N}$  such that k > m,

$$\sum_{i=1}^{n} \frac{a_i^k}{b_i^m} \ge n^{1+m-k} \frac{\left(\sum_{i=1}^{n} a_i\right)^k}{\left(\sum_{i=1}^{n} b_i\right)^m}.$$

#### Variation: Generalized Titu's Lemma

For any real numbers  $a_1, ..., a_n$ , positive real numbers  $b_1, ..., b_n$ , positive integers  $m, k \in \mathbb{N}$  such that k > m,

$$\sum_{i=1}^{n} \frac{a_i^k}{b_i^m} \ge n^{1+m-k} \frac{\left(\sum_{i=1}^{n} a_i\right)^k}{\left(\sum_{i=1}^{n} b_i\right)^m}.$$

*Proof.* Focus on the "draft matrix"

$$\begin{bmatrix} a_1^k/b_1^m & b_1 & \cdots & b_1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_n^k/b_n^m & b_n & \cdots & b_n & 1 & \cdots & 1 \end{bmatrix}$$

with m columns of b's and k-1-m columns of 1's.

#### Last example

Most other questions are direct consequences of the previous lemma, including the following:

Exercise. For any a, b, c > 0 satisfying abc = 1, and positive  $k \ge 2$ , show that

$$\sum_{\text{cyc}} \frac{1}{a^k (b+c)} \ge \frac{3}{2}.$$

#### Last example

Most other questions are direct consequences of the previous lemma, including the following:

*Exercise.* For any a, b, c > 0 satisfying abc = 1, and positive  $k \ge 2$ , show that

$$\sum_{\text{cyc}} \frac{1}{a^k (b+c)} \ge \frac{3}{2}.$$

Attempt: Replace the numerator on LHS by abc. Then

$$\text{LHS} = \sum_{\text{cyc}} \frac{\left(\frac{1}{a}\right)^{k-1}}{\frac{1}{b} + \frac{1}{c}} \ge \frac{\left(\sum_{\text{cyc}} \frac{1}{a}\right)^{k-1}}{2\sum_{\text{cyc}} \frac{1}{a}} = \frac{1}{2} \cdot 3^{1+1-(k-1)} \cdot \left(\sum_{\text{cyc}} \frac{1}{a}\right)^{k-2} \ge \frac{1}{2} \cdot 3^{(3-k)+(k-2)}$$

*Problem*: To apply the generalized Cauchy–Schwarz inequality, we need k-1>1, i.e. k>2. The author of the <u>original article</u> doesn't address the case when k=2,

#### Last example (continued)

which turns out to be the Nesbitt's inequality: For any positive real numbers a, b, c, we have

$$\sum_{\text{cyc}} \frac{a}{b+c} \ge \frac{3}{2}.$$

Observe that the above inequality is homogeneous, so WLOG, we can assume a+b+c=1. Then it's equivalent to

$$\sum_{\text{cyc}} \frac{a+b+c}{b+c} \ge \frac{3}{2} + 3.$$

The numerator on LHS is  $1 = 1^2$ , so that <u>Titu's Lemma</u> can be used.