Symmetric means and Muirhead's Inequality ①

We considered the inequality

$$a+b+c \leq \frac{a^3}{bc}+\frac{b^3}{ac}+\frac{c^3}{ab}$$

in the exercises. We now are going to understand this in a more general way.

Definition. Let $\sum! F(a_1,\dots,a_n)$ be the sum of the $n!$ terms given by ~~ex~~ evaluating F on all permutations of $a_1,\dots,a_n$.

~~Def~~

Given $(\alpha_1,\dots,\alpha_n)$, we define the <u>symmetrical mean</u> $[\alpha](\vec{a})$ by $\frac{1}{n!}\sum! a_1^{\alpha_1}\cdots a_n^{\alpha_n}$.

Examples. $[1,0,\dots,0] = \frac{(n-1)!}{n!}(a_1+\dots+a_n) = \mathcal{U}(\vec{a})$

$[1/n,\dots,1/n] = \frac{n!}{n!}(a_1^{1/n}\cdots a_n^{1/n}) = G(\vec{a})$

$$[3,-1,-1] = \frac{1}{3!}\left(2a^{3}b^{-1}c^{-1} + 2a^{-1}b^{3}c^{-1} + 2a^{-1}b^{-1}c^{3}\right.$$

We want to figure out when $[\alpha]$ is comparable to $[\alpha']$.

Definition. We say that $\vec{\alpha}'$ is majorised by $\vec{\alpha}$ (as vectors) when we can permute their entries so

1) $\alpha_1' + \dots + \alpha_n' = \alpha_1 + \dots + \alpha_n$

2) $\alpha_1' \geq \alpha_2' \geq \dots \geq \alpha_n'$, $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$

3) $\alpha_1' \leq \alpha_1$, $\alpha_1' + \alpha_2' \leq \alpha_1 + \alpha_2, \dots, \alpha_1' + \dots + \alpha_n' \leq \alpha_1 + \dots + \alpha_n$

We denote this by $\alpha' \prec \alpha$.

Muirhead's Theorem. The symmetric means $[\alpha']$ and $[\alpha]$ are comparable $\Leftrightarrow$ one of α', α majorizes the other. If $\alpha' \prec \alpha$, $[\alpha'] \leq [\alpha]$ with equality only if $\alpha' = \alpha$ or all a_i are equal.

Proof. Suppose $[\alpha']$ and $[\alpha]$ are comparable, and wlog that $[\alpha'] \leq [\alpha]$ for all $\vec{a}$.

If $\vec{a} = (x, x, \dots, x)$ then

$$[\alpha'] = \frac{1}{n!} \sum ! \, x^{\sum_i \alpha_i'} = x^{\sum_i \alpha_i'}$$

$$[\alpha] = x^{\sum_i \alpha_i}$$

If $x = e$ $\quad e^{\sum_i \alpha_i'} \leq e^{\sum_i \alpha_i} \overset{\text{log of both sides}}{\Rightarrow} \sum_i \alpha_i' \leq \sum_i \alpha_i$

(because log is an increasing function).

If $x = e^{-1}$, $\quad e^{-\sum_i \alpha_i'} \leq e^{-\sum_i \alpha_i} \Rightarrow -\sum_i \alpha_i' \leq -\sum_i \alpha_i$

$\Rightarrow \sum_i \alpha_i' \geq \sum_i \alpha_i$. Thus we must have

$\sum_i \alpha_i' = \sum_i \alpha_i$. (Condition 1)

~~log~~

Now suppose we let $a_1 = \dots = a_k = x$ and $a_{k+1} = \dots = a_n = 1$. ~~Then~~ Further, wlog we

assume that $\alpha_1' \geq \alpha_2' \geq \ldots \geq \alpha_n'$ and
$\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n$. (condition 2) Now

$[\alpha'] =$ a polynomial in x with powers given by all possible sums of K of the α_i'.

$$= x^{\alpha_1' + \ldots + \alpha_K'} + \text{lower order terms}$$

Similarly

$$[\alpha] = x^{\alpha_1 + \ldots + \alpha_K} + \text{lower order terms.}$$

We can have $[\alpha'] \leq [\alpha]$ for all $x \to \infty$ only if $\alpha_1' + \ldots + \alpha_K' \leq \alpha_1 + \ldots + \alpha_K$, establishing that we have condition 3 and $\alpha' \prec \alpha$.

The proof of sufficiency is harder.

Definition. We define a linear transformation of type T ~~between~~ ~~α_k~~ of $\vec{\alpha} = (\alpha_1, \ldots, \alpha_n)$ as follows.

Let α_k, α_ℓ be any two unequal elements of $\vec{\alpha}$, with $\alpha_k > \alpha_\ell$. We can ~~pick~~ let $\rho = \frac{\alpha_k + \alpha_\ell}{2}$ ~~between α_k and α_ℓ (example: $\rho = \frac{\alpha_k+\alpha_\ell}{2}$)~~ ~~so that~~ and write

$$\alpha_k = \rho + \gamma \qquad \alpha_\ell = \rho - \gamma$$

where $\gamma = \frac{\alpha_k - \alpha_\ell}{2}$. Clearly, $0 \leq \gamma \leq \rho$.

For any $0 \leq \sigma < \gamma$, we can define

$$(T\vec{\alpha})_i = \begin{cases} \rho + \sigma, & \text{if } i = k \\ \rho - \sigma, & \text{if } i = \ell \\ \alpha_i, & \text{otherwise} \end{cases}$$

Lemma. If $\alpha' = T\alpha$, then $[\alpha'] \leq [\alpha]$, with equality only if all a are equal.

Proof. Wlog, assume $k=1$, $\ell=2$ (since $[\alpha]$ is symmetric in the entries of $\vec{\alpha}$, rearranging $\vec{\alpha}$ doesn't change the value).

$$n!\,2[\alpha] - n!\,2[\alpha'] = n!\,2[\rho+\gamma, \rho-\gamma, \alpha_3, \ldots, \alpha_n] - n!\,2[\rho+\sigma, \rho-\sigma, \alpha_3, \ldots, \alpha_n]$$

$$= \sum! \, a_3^{\alpha_3} \cdots a_n^{\alpha_n} \left(a_1^{\rho+\gamma} a_2^{\rho-\gamma} + a_1^{\rho-\gamma} a_2^{\rho+\gamma} - a_1^{\rho+\sigma} a_2^{\rho-\sigma} - a_1^{\rho-\sigma} a_2^{\rho+\sigma} \right)$$

$$= \sum! \, (a_1 a_2)^{\rho-\gamma} a_3^{\alpha_3} \cdots a_n^{\alpha_n} \Big(a_1^{\rho+\gamma-(\rho-\gamma)} a_2^{\rho-\gamma-(\rho-\gamma)} + a_1^{\rho-\gamma-(\rho-\gamma)} a_2^{\rho+\gamma-(\rho-\gamma)} - a_1^{\rho+\sigma-(\rho-\gamma)} a_2^{\rho-\sigma-(\rho-\gamma)} - a_1^{\rho-\sigma-(\rho-\gamma)} a_2^{\rho+\sigma-(\rho-\gamma)} \Big)$$

$$= \sum ! (a_1 a_2)^{\rho-\gamma} a_3^{\alpha_3} \cdots a_n^{\alpha_n} \left(a_1^{2\gamma} + a_2^{2\gamma} - a_1^{\sigma+\gamma} a_2^{\gamma-\sigma} - a_1^{\gamma-\sigma} a_2^{\sigma+\gamma} \right)$$

$$= \sum ! (a_1 a_2)^{\rho-\gamma} a_3^{\alpha_3} \cdots a_n^{\alpha_n} \left(a_1^{\gamma+\sigma} - a_2^{\gamma+\sigma} \right) \left(a_1^{\gamma-\sigma} - a_2^{\gamma-\sigma} \right)$$

Now $\gamma - \sigma > 0$ and $\gamma + \sigma > 0$, so ~~since~~ the last two terms have the same sign whether $a_1 \geq a_2$ or $a_1 \leq a_2$ in this permotation. Thus all terms are ≥ 0, with equality only if $a_1 = a_2$ in every permutation $\Rightarrow$ all a_i are equal. $\square$

Lemma 2. If $\vec{\alpha}' \prec \vec{\alpha}$, but $\vec{\alpha}' \neq \vec{\alpha}$ then $\vec{\alpha}'$ can be obtained from $\vec{\alpha}$ by a finite number of transformations of type T.

Proof. ~~Suppose that~~ Since $\vec{\alpha}' \prec \vec{\alpha}$, we know

$$\alpha_1' \geq \alpha_2' \geq \cdots \geq \alpha_n' \quad \text{and} \quad \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$$

and $\alpha_1' + \dots + \alpha_n' = \alpha_1 + \dots + \alpha_n$.

Call the number of indices k for which $\alpha_k' \neq \alpha_k$ the discrepancy, r. We will show that if the lemma is true for discrepancy $< r$, then it is true for discrepancy r. If $r=0$ then $\vec{\alpha}' = \vec{\alpha}$ and there is nothing to prove. So suppose $r>0$.

We know

$$\sum_k \alpha_k - \alpha_k' = 0$$

and not all $\alpha_k - \alpha_k'$ are zero. So some must be positive and others negative. If k^* is the first difference, since

$$\underbrace{\alpha_1' + \dots + \alpha_{k^*-1}'} + \alpha_{k^*}' \le \underbrace{\alpha_1 + \dots + \alpha_{k^*-1}} + \alpha_{k^*}$$

equal

we know $\alpha_{k^*} - \alpha_{k^*}' > 0$ (if $=0$, not a difference!)

This means that eventually we can find a positive difference followed by a negative difference (maybe with some equal terms in between) of some k, ℓ so that

$$\alpha'_k < \alpha_k, \quad \underbrace{\alpha'_{k+1} = \alpha_{k+1}, \ldots, \alpha'_{\ell-1} = \alpha_{\ell-1}}_{\text{these may be absent altogether}}, \quad \alpha'_\ell > \alpha_\ell$$

and $k < \ell$. We take $\alpha_k = \rho + \gamma$, $\alpha_\ell = \rho - \gamma$ and let $\sigma = \mathrm{Max}(|\alpha'_k - \rho|, |\alpha'_\ell - \rho|)$

(Idea: We're going to fix the larger difference, but we don't know which one that is.)

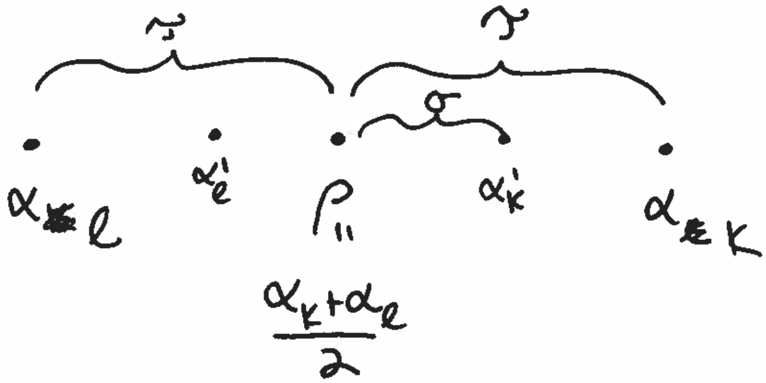
Now $\alpha_k > \alpha_\ell$ (proof: $\alpha_k > \alpha'_k \geq \alpha'_\ell > \alpha_\ell$), so $0 < \gamma \leq \rho$ as before ($\rho = \frac{\alpha_k + \alpha_\ell}{2}$, $\gamma = \frac{\alpha_k - \alpha_\ell}{2}$).

Further either (we claim!)

$$\alpha'_\ell - \rho = -\sigma \quad \text{or} \quad \alpha'_k - \rho = \sigma$$

~~(Proof: Suppose~~ $\cancel{\sigma = |\alpha'_\ell - \rho| \geq |\alpha'_k - \rho|}$. ~~We need to~~

Now we know $\alpha_k > \alpha_k' \geq \alpha_\ell' > \alpha_\ell$, so the picture is



and wherever α_k' and α_ℓ' are in $[\alpha_\ell, \alpha_k]$, they are certainly no farther from ρ than τ.
Thus $0 \leq \sigma < \tau \leq \rho$.

Further, looking at the picture, if $\alpha_\ell' > \rho$ then $\alpha_k' - \rho \geq \alpha_\ell' - \rho > 0$ and $\sigma = \alpha_k' - \rho$. If $\alpha_k' < \rho$ then $\rho - \alpha_\ell' \geq \rho - \alpha_k' > 0$, so $\sigma = \rho - \alpha_\ell'$ or $\alpha_\ell' - \rho = -\sigma$. If $\alpha_\ell' < \rho < \alpha_k'$ then either $\alpha_k' - \rho = \sigma$ or $\alpha_\ell' - \rho = -\sigma$ (depending on whether $|\alpha_k' - \rho| > |\alpha_\ell' - \rho|$
We now let

$$\alpha_k'' = \rho + \sigma \qquad \alpha_\ell'' = \rho - \sigma, \quad \alpha_i'' = \alpha_i \text{ (for all other indices } i\text{)}$$

Clearly, $\alpha'' = T\alpha$. Now ~~#~~ either

1) $\alpha_k' - \rho = \sigma$, so $\alpha_k'' = \rho + \sigma = \rho + \alpha_k' - \rho = \alpha_k'$

or

2) $\alpha_\ell' - \rho = -\sigma$, so $\alpha_\ell'' = \rho - \sigma = \rho + \alpha_\ell' - \rho = \alpha_\ell'$

(or both), so the discrepancy between $\vec{\alpha}''$ and $\vec{\alpha}$ is lower (either $r-1$ or $r-2$).

We now have to show that $\vec{\alpha}'' \succ \vec{\alpha}'$ (then we can use the inductive step to assume that we can get from $\vec{\alpha}''$ to $\vec{\alpha}$ by more transformations of type T).

We must check 3 conditions.

1) $\alpha_k'' + \alpha_\ell'' = \rho + \sigma + \rho - \sigma = 2\rho = \rho + \gamma + \rho - \gamma = \alpha_k + \alpha_\ell$

So $\sum \alpha_i'' = \sum \alpha_i$. And $\sum \alpha_i = \sum \alpha_i'$ (b/c $\vec{\alpha} \succ \vec{\alpha}'$)

so $\sum \alpha_i'' = \sum \alpha_i'$, as required.

2) We have only changed α''_K and α''_ℓ, so we have to check

$$\alpha''_{K-1} = \alpha_{K-1} \geq \alpha_K = \rho + \gamma > \rho + \sigma = \alpha''_K$$

Now

$$\alpha''_K = \rho + \sigma \geq \rho + \overbrace{|\alpha'_K - \rho|} \geq \rho + \overbrace{\alpha'_K - \rho}^{\text{If this is negative, it's certainly } \leq |\alpha'_K - \rho|} = \alpha'_K$$

$\uparrow$ remember $\sigma = \mathrm{Max}(|\alpha'_K - \rho|, |\alpha'_\ell - \rho|)$

so

$$\alpha''_K \geq \alpha'_K \geq \underbrace{\alpha'_{K+1} = \alpha_{K+1}}_{\text{these are part of the "middle run" of equalities}} = \alpha''_{K+1}$$

Now

$$|\rho - \alpha'_\ell| + |\alpha'_\ell| \geq |\rho| \geq \rho - \alpha'_\ell$$

$$\alpha''_\ell = \rho - \sigma \leq \rho - |\alpha'_\ell - \rho| \leq \rho - (\rho - \alpha'_\ell) = \alpha'_\ell$$

so

$$\alpha''_{\ell-1} = \alpha_{\ell-1} = \alpha'_{\ell-1} \geq \alpha'_\ell \geq \alpha''_\ell$$

and

$$\alpha''_\ell = \rho - \sigma > \rho - \gamma = \alpha_\ell \geq \alpha_{\ell+1} = \alpha''_{\ell+1}.$$

So we've checked the α''_i are still in order.

3) We last need to check the partial sum conditions.

$$\alpha'_1 + \alpha'_2 + \ldots + \alpha'_i \leq \alpha''_1 + \alpha''_2 + \ldots + \alpha''_i.$$

If $i<k$, this is true because all the α'' terms are equal to α terms (and $\alpha' \prec \alpha$).

If $i=k$, we recall $\alpha'_k \leq \alpha''_k$ (from above). ~~If $k<i<\ell$~~, so this is true b/c it is true for ~~$i<$~~ $i=k-1$.

If $\ell > i > k$, we ~~already~~ have

equal pairwise

$$\underbrace{\alpha'_1 + \ldots + \alpha'_k}\; + \overbrace{\alpha'_{k+1} + \ldots + \alpha'_i} \leq \underbrace{\alpha''_1 + \ldots + \alpha''_k}\; + \overbrace{\alpha''_{k+1} + \ldots + \alpha''_i}$$

proved $\leq$ above

Finally, if $i \geq \ell$,

$$\underbrace{\alpha'_1 + \dots + \alpha'_\ell}_{} + \dots + \alpha'_i \overset{\text{by } \alpha' \prec \alpha}{\leq} \underbrace{\alpha_1 + \dots + \alpha_\ell}_{} + \dots + \alpha_i$$

$$= \underbrace{\alpha''_1 + \dots + \alpha''_\ell}_{\text{by above}} + \underbrace{\dots + \alpha''_i}_{\text{these terms of } \alpha'' \text{ are equal to } \alpha}$$

$$\leq \alpha''_1 + \dots + \alpha''_i$$

This proves $\alpha'' \succ \alpha$, as required. $\square$

We have now proved Muirhead's Theorem!

Alternative form.

Suppose P is an $n \times n$ matrix with all rows and columns summing to 1. Then P is called "doubly-stochastic".

Proposition. $\vec{\alpha}' \prec \vec{\alpha}$ iff $\exists$ a doubly-stochastic P so that $\vec{\alpha}' = P\vec{\alpha}$.

Corollary. If $\alpha_1 + \dots + \alpha_n = 1$ then

$$G(\vec{a}) \leq [\alpha](\vec{a}) \leq \mathcal{U}(\vec{a})$$

with equality only if $[\alpha] = G$, $[\alpha] = \mathcal{U}$ or all a_i are equal.

Proof. Recall $G = [(1/n, \dots, 1/n)]$, $\mathcal{U} = [(1, 0, \dots, 0)]$.

Now

$$\frac{1}{n} = \frac{\alpha_1}{n} + \dots + \frac{\alpha_n}{n}$$

so $(1/n, \dots, 1/n)^T = \begin{bmatrix} 1/n & \cdots & 1/n \\ \vdots & & \\ 1/n & \cdots & 1/n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = P\alpha$, and

$$(1/n, \dots, 1/n) \prec \alpha.$$

Further

$$(\alpha_1, \dots, \alpha_n)^T = \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \\ \alpha_2 \alpha_3 & \cdots & \alpha_1 \\ \vdots & & \\ \alpha_n \alpha_1 & \cdots & \alpha_{n-1} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = P \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

so

$$\alpha \prec (1, 0, \dots, 0).$$

□