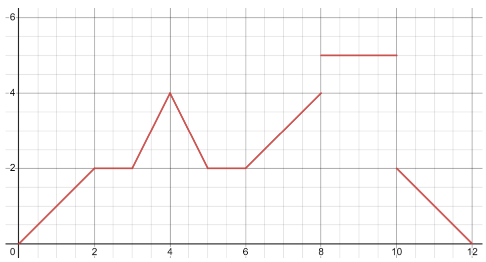
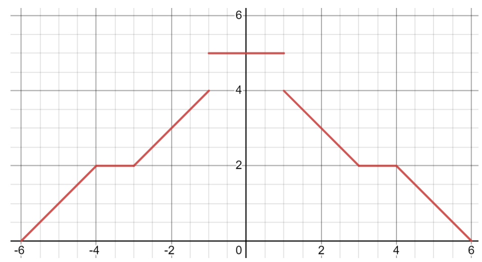
What is the Riesz Rearrangement Inequality?

Symmetric Decreasing Rearrangements

Given a function $f\colon \mathbb{R}^n \to [0,\infty)$, we want to define a new function $f^*\colon \mathbb{R}^n \to [0,\infty)$ with the same distribution of values, but such that the largest values are maximally concentrated near the origin. For example, consider the following functions:





One can think of taking small columns of mass under the graph on the left and re-shuffling them to get the graph on the right, so that the largest columns are near the center. Note that at each height, the measure of points with values above that height is the same between the two graphs.

The most obvious first implementation of this is for (characteristic functions of) sets. Given a finite-measure $A \subset \mathbb{R}^n$, define

$$A^* = \{x \,:\, \omega_n \cdot |x|^n < \mathcal{L}(A)\},$$

where $\mathcal{L}$ is the n-dimensional Lebesgue measure, and ω_n is the volume of the unit n-ball. The result is that $\mathcal{L}(A^*) = \mathcal{L}(A)$. To extend this definition to functions, we have the following:

Definition. If $f\colon \mathbb{R}^n \to [0,\infty)$ is a measurable function, we define the *symmetric decreasing rearrangement* $f^*\colon \mathbb{R}^n \to \mathbb{R}$ by $f^*(x) = \int_0^\infty \chi_{\{f>t\}^*}(x)\,dt$, assuming each super-level set $\{f > t\}$ has finite measure (in which case we say f *vanishes at infinity*). We additionally require that f is essentially bounded, so that $f^*(0) < \infty$.

This definition essentially asks: “In the rearrangement, which output values should have layers big enough to include x?” The integral then measures the length of the interval of such values. Some equivalent definitions use an infimum or supremum to find the length of the interval.

The definition produces the appropriate properties:

- f^* is radially symmetric, since each $\{f > t\}^*$ is.
- f^* is decreasing with respect to norm: if $|x| \leqslant |y|$, then $\omega_n \cdot |y|^n < \mathcal{L}(\{|f| > t\})$ implies $\omega_n \cdot |x|^n < \mathcal{L}(\{|f| > t\})$, so $\chi_{\{|f|>t\}}(y) \leqslant \chi_{\{|f|>t\}}(x)$ for each t, and so $f^*(y) \leqslant f^*(x)$.
- f^* is a rearrangement: we have $\mathcal{L}(\{f > t\}) = \mathcal{L}(\{f^* > t\})$ for all t, i.e., the two functions have the same distribution. A bit stronger, we have $\{f^* > t\} = \{f > t\}^*$. To show this, we

have

$$\begin{aligned} & x \in \{f^* > t\} \\ \iff & \int_{\{f>s\}*} (x)\,ds > t \\ \iff & (\exists s > t)\ \chi_{\{f>s\}*}(x) = 1 \\ \iff & (\exists s > t)\ \omega_n |x|^n < \mathcal{L}(\{f > s\}) \end{aligned} \qquad \begin{aligned} & x \in \{f > t\}^* \\ \iff & \omega_n |x|^n < \mathcal{L}(\{f > t\}) \end{aligned}$$

The left-hand statement trivially implies the right, and the right-hand statement implies the left by the right-continuity of distribution functions.

The Reisz Rearrangement Inequality will state that these symmetric decreasing rearrangements solve the optimization problem of maximizing a certain kind of integral.

A first inequality

We begin with a simpler inequality to demonstrate working with rearrangements.

Theorem. *If $f, g\colon \mathbb{R}^n \to [0, \infty)$ are measurable, essentially bounded, and vanish at infinity, then $\int fg \leqslant \int f^* g^*$.*

This can be seen as a generalization of the following elementary statement: if $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ is a permutation, and $a_1 \leqslant \cdots \leqslant a_n$ and $b_1 \leqslant \ldots \leqslant b_n$ are increasing lists of real numbers, then $\sum a_i b_{\sigma(i)} \leqslant \sum a_i b_i$. In other words, to maximize the value of a sum of products, one should pair large values with large values and small values with small values.

Proof. We can re-write the left-hand side to apply Fubini's theorem:

$$\begin{aligned} \int_{\mathbb{R}^n} f(x) g(x)\,dx &= \int_{\mathbb{R}^n} \left(\int_0^\infty \chi_{\{f>t\}}(x)\,dt \right) \left(\int_0^\infty \chi_{\{g>s\}}(x)\,ds \right) dx \\ &= \int_0^\infty \int_0^\infty \int_{\mathbb{R}^n} \chi_{\{f>t\}}(x) \chi_{\{g>s\}}(x)\,dx\,ds\,dt \end{aligned}$$

We can expand the definitions of the right-hand side and similarly apply Fubini;

$$\begin{aligned} \int_{\mathbb{R}^n} f^*(x) g^*(x)\,dx &= \int_{\mathbb{R}^n} \left(\int_0^\infty \chi_{\{f>t\}^*}(x)\,dt \right) \left(\int_0^\infty \chi_{\{g>s\}^*}(x)\,ds \right) dx \\ &= \int_0^\infty \int_0^\infty \int_{\mathbb{R}^n} \chi_{\{f>t\}^*}(x) \chi_{\{g>s\}^*}(x)\,dx\,ds\,dt. \end{aligned}$$

It thus suffices to show that for each $t, s > 0$ we have

$$\int_{\mathbb{R}^n} \chi_{\{f>t\}}(x) \chi_{\{g>s\}}(x)\,dx \leqslant \int_{\mathbb{R}^n} \chi_{\{f>t\}^*}(x) \chi_{\{g>s\}^*}(x)\,dx$$

Because $\chi_{A^*} = (\chi_A)^*$, for any measurable set A, we have reduced the situation to assuming that f and g are characteristic functions of sets. Indeed, if A and B are measurable, we shall show that $\int \chi_A \chi_B \leqslant \int \chi_A^* \chi_B^*$, i.e., $\mathcal{L}(A \cap B) \leqslant \mathcal{L}(A^* \cap B^*)$. Assuming $\mathcal{L}(A) \leqslant \mathcal{L}(B)$, we have $A^* \subset B^*$, and so $\mathcal{L}(A^* \cap B^*) = \mathcal{L}(A^*) = \mathcal{L}(A) \geqslant \mathcal{L}(A \cap B)$. $\square$

Corollary. *For any f, g where f^* and g^* are defined,* $\|f^* - g^*\|_2 \leqslant \|f - g\|_2$.

Proof.

$$\begin{aligned} \|f^* - g^*\|_2^2 - \|f - g\|_2^2 &= \int_{\mathbb{R}^n} (f^2 - 2fg + g^2) - \int_{\mathbb{R}^n} (f^{*2} - 2f^*g^* + g^{*2}) \\ &= \underbrace{\|f\|_2^2 - \|f^*\|_2^2}_{0} + \underbrace{\|g\|_2^2 - \|g^*\|_2^2}_{0} + 2\left(\int f^*g^* - \int fg\right) \geqslant 0. \end{aligned}$$

□

RRI statement

This is Riesz Rearrangement Inequality:

Theorem. *If $f, g, h \colon \mathbb{R}^n \to [0, \infty)$ are measurable, essentially bounded, and vanish at infinity, then*

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x)g(x-y)h(y)\,dx\,dy \leqslant \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f^*(x)g^*(x-y)h^*(y)\,dx\,dy.$$

Unlike the simpler inequality from the previous section, this depends more on the structure of $\mathbb{R}^n$. In the simpler inequality, it was only important that large values be rearranged so as to be multiplied by large values, but where exactly those values occurred did not matter (any measure-preserving transformation could be applied to $\mathbb{R}^n$ and preserve $\int fg$).

On the other hand, when maximizing the three-function inequality, the inclusion of the $g(x-y)$ factor encourages not only that large values of f be paired with large values of h, but also that the large values of each function be centralized in one area.

We may also write $I(f, g, h) := \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x)g(x-y)h(y)\,dx\,dy$, so the inequality becomes $I(f, g, h) \leqslant I(f^*, g^*, h^*)$.

Proof in one dimension

Assume $n = 1$.

Proof. By the same Fubini trick as the previous proof, it suffices to only consider finite-measure sets $F, G, H \subset \mathbb{R}$ and then show that

$$I(\chi_F, \chi_G, \chi_H) \leqslant I(\chi_{F^*}, \chi_{G^*}, \chi_{H^*})$$

Then by the outer-regularity of Lebesgue measure, the sets can be approximated by decreasing sequences of open sets, (F_k), (G_k), (H_k). By the dominated convergence theorem, $I(\chi_{F_k}, \chi_{G_k}, \chi_{H_k}) \to I(\chi_F, \chi_G, \chi_H)$, and certainly $I(\chi_{F_k^*}, \chi_{G_k^*}, \chi_{H_k^*}) \to I(\chi_{F^*}, \chi_{G^*}, \chi_{H^*})$. Showing the limit of each term in the sequence this shows the final inequality in the limit, so it suffices to assume that F, G, and H are open.

Open sets F of $\mathbb{R}$ are countable unions of disjoint intervals, since we can take all $[k, k+1) \subset F$, $k \in \mathbb{Z}$, followed by all remaining $[k/2, (k+1)/2)$, and continue shrinking the acceptable interval size by a factor of $1/2$.

We can thus approximate F, G, and H by finite unions of intervals: Let F_k be the union of the first k of the countable many intervals. By monotone convergence, $I(\chi_{F_m}, \chi_{G_m}, \chi_{H_m}) \to I(\chi_F, \chi_G, \chi_H)$ and $I(\chi_{F_m^*}, \chi_{G_m^*}, \chi_{H_m^*}) \to I(\chi_{F^*}, \chi_{G^*}, \chi_{H^*})$ Hence, by showing the bound at each finite step, we show the desired bound in the limit. We may thus assume that F, G, and H are finite disjoint unions of intervals. Write

$$f(x) = \chi_F(x) = \sum f_i(x - a_i), \quad g(x) = \chi_G(x) = \sum g_j(x - b_j), \quad h(x) = \chi_H(x) = \sum h_k(x - c_k),$$

whee f_i, g_j, h_k are characteristic functions of intervals centered at 0. Then define

$$I_t := \sum_{i,j,k} \int_{\mathbb{R}} \int_{\mathbb{R}} f_i(x - ta_i)\, g_j\big((x-y) - tb_j\big)\, h_k(y - tc_k)\, dx\, dy,$$

so that $I_1 = I(f, g, h)$. Changing variables gives that

$$I_t = \sum_{i,j,k} \int_{\mathbb{R}} \int_{\mathbb{R}} f_i(x)\, g_j(x - y)\, h_k\big(y + (a - b - c)t\big)\, dx\, dy$$

By defining $u_{ij}(y) := \int_{\mathbb{R}} f_i(x) g_j(x - y)\, dx$, we have

$$I_t = \sum_{i,j,k} \int_{\mathbb{R}} u_{ij}(y)\, h_k\big(y + (a - b - c)t\big)\, dy$$

Now suppose we begin to vary the parameter t from 1 down toward 0. The graph of u_{ij} is shaped like a trapeziod centered at 0, and as we slide h_k past it, the integral is maximized when h_k is centered, i.e., when $t = 0$. As we slowly decrease t, I_t can only increase. When any two intervals defining one of the functions f, g, h touch, we can replace them by one larger interval. After finitely many steps, we have that f, g, and h are characteristic functions of balls centered at the origin, so these are the symmetric decreasing rearrangements of the original functions, and $I(f, g, h)$ has not decreased. □

About the proof in higher dimensions

- The higher-dimensional proof uses the *Steiner symmetrization*, where along each line parallel to some direction, the values along that line are symmetrically rearranged, which does not increase the total value of the integral. When rearranging entire hyperplanes instead of just lines, this is called the *Schwarz symmetrization.*

- By repeatedly symmetrizing along different axes and applying a compactness argument about the sphere S^{n-1} of possible directions, we can achieve the symmetric decreasing rearrangements of the functions, without having increased the integral's value.

Extra facts

- If g is *strictly* symmetric-decreasing, then $I(f,g,h) = I(f^*,g^*,h^*)$ if and only if f and h are almost everywhere equal to shifts of symmetric decreasing functions. For all other f and h, the inequality is strict.
- There are many more generalizations involving products of more functions with more linear combinations of variables as arguments.
- The RRI was used in solving several kinds of optimization problems, including the isoperimetric inequality (that balls minimize surface area), the Pólya-Szegö inequality about Sobolev energy, and the fact that balls minimize electrostatic capacity in physics.
- Most of the material in this talk was from Lieb and Loss's 2001 graduate *Analysis* textbook, where it was included for its importance, as well as because "From the pedagogic point of view, it provides a good exercise in manipulating measurable sets."
- The original proof was given by Riesz in 1930.

References

- Lieb, Elliott; Loss, Michael (2001). *Analysis.* Graduate studies in Mathematics. Vol 14 (2nd ed.). American Mathematical Society. ISBN 978-0821827833.
- Riesz, Frigyes (1930). "Sur une inégalité intégrale". Journal of the London Mathematical Society. 5 (3): 162–168. doi:10.1112/jlms/s1-5.3.162