An optimization approach to the Riemann Hypothesis

By: Hassine Saidane, PhD - UIUC

Abstract

The optimization of theoretical concepts such as action or utility functions enabled the derivation of important theories and laws in some scientific fields such as physics and economics. These breakthroughs suggested that the problem of the location of the Riemann Zeta Function's (RZF) nontrivial zeros can be similarly addressed in a mathematical programming framework. Using a constrained nonlinear optimization formulation of the problem, we prove that RZF's nontrivial zeros are located on the critical line, thereby confirming the Riemann Hypothesis. This result is a direct implication of the Karush-Kuhn-Tucker optimality conditions associated with the formulated nonlinear program.

Keywords
Riemann Zeta function, Riemann Hypothesis, Constrained Optimization, Karush-Kuhn-Tucker conditions.

Introduction

A great deal of research has been and still is being devoted to the zeros of the Riemann Zeta function (RZF) that are located in the critical strip[1] and known as the nontrivial zeros of RZF. The Riemann Hypothesis (RH) states that these zeros are located on the critical line[2]. Although a large number of nontrivial zeros have proved to be located on the critical line through numerical computation methods, starting with Riemann's manual computation of the first few zeros [1], no analytical proof or disproof of RH has been developed since its conjecture by Riemann in 1859.

In this paper, we propose an analytical approach to RH based on optimization. This tool proved successful in deriving some important scientific theories and laws [2]. By formulating and solving the appropriate optimization problem, we derive evidence in support of the Riemann Hypothesis.

Problem formulation

We denote the Riemann zeta function (RZF) as $\zeta(\sigma+it) = U(\sigma,t)+iV(\sigma,t)$, for complex $s = \sigma+it$. As a consequence of the properties of RZF and the properties of its nontrivial zeros[3], the search for the location of these zeros can be limited to the left half of the critical strip since zeros on the right of the critical line can be obtained by symmetry about this line as per RZF's property (6)[4]. Also RZF's functional equation shows that nontrivial zeros occur either in singles on the critical line, or in pairs, off of the critical that are symmetric about this line.

---

[1] *A strip in the complex plane defined by $0 \le \sigma \le 1$*
[2] *The critical line is the line in the complex plane defined by $\sigma = 1/2$*
[3] *See Appendix for the list of properties*
[4] *See Appendix for the list of properties*

Hence, this search entails finding the value σ^* where $\zeta(\sigma^*+it)$ or equivalently $|\zeta(\sigma^*;\, t)|^2$, vanishes at some height $t = t^*$. In this framework, this task can be accomplished by minimizing the simple objective function $|\zeta(\sigma;\, t^*)|^2$ under the constraint $0 \le \sigma \le ½$, with t^* being a constant.

The optimization problem of interest is then to:

$$\begin{aligned} &\text{Minimize } f(\sigma) = Z(\sigma;\, t^*) = |\zeta(\sigma;\, t^*)|^2 = U^2(\sigma;\, t^*) + V^2(\sigma;\, t^*) \\ &\text{Subject to: } g(\sigma) = \sigma - 1/2 \le 0 \qquad\qquad \text{(P)} \\ &\qquad\qquad \sigma \ge 0 \end{aligned}$$

With $f(\sigma)$ and $g(\sigma)$ twice differentiable for σ in [0,1/2] .

To solve the nonlinear constrained problem (P), we use the Karush-Kuhn-Tucker (KKT) method [3] with a nonnegativity condition on the variable σ. The Lagrange function associated with (P) is then:

$$L(\sigma, \mu;\, t^*) = Z(\sigma;\, t^*) + \mu(\sigma - ½)$$

Where μ *is* the Lagrange multipliers associated with the constraint $g(\sigma)$.

For minimization problems such as problem (P), with continuously differentiable functions, nonnegative variables (σ in our case), and under a regularity qualification of the constraints[5], optimality requires the existence of a vector $v^* = (\sigma^*, \mu^*)$ that meets the necessary KKT conditions with nonnegative variables [4]. Using the notation below:

$$L_\sigma = \partial L/\partial\sigma,\; L_\mu = \partial L/\partial\mu,\; Z_\sigma = \partial Z/\partial\sigma,\; L_{\sigma\sigma} = \partial^2 L/\partial\sigma^2,\; L_{\mu\mu} = \partial^2 L/\partial\mu^2$$

The necessary KKT conditions are:

1. Complementary slackness conditions

$$\sigma^* L_\sigma(\sigma^*, \mu^*;\, t^*) = \sigma^*(Z_\sigma^* + \mu^*) = 0 \qquad (1)$$
$$\mu^* L_\mu(\sigma^*, \mu^*;\, t^*) = \mu^*(\sigma^* - ½) = 0 \qquad (2)$$

2. Feasibility conditions

$$L_\sigma(\sigma^*, \mu^*;\, t^*) = Z_\sigma^* + \mu^* \ge 0 \qquad (3)$$
$$L_\mu(\sigma^*, \mu^*;\, t^*) = (\sigma^* - ½) \le 0 \qquad (4)$$
$$\mu^* \ge 0 \qquad (5)$$
$$\sigma^* \ge 0 \qquad (6)$$

Solving the system of KKT necessary conditions will enable the identification of candidate solutions to optimization problems such as problem (P). To do this, it is necessary to consider all the subsets of systems of equations defined by the complementarity conditions using the

[5] *The gradients of the equality and binding nonequality constraints have to be linearly independent at the stationary/critical point(s) of the Lagrangian function. This requirement is of no concern here since we have one constraint only.*

various combinations of the factors involved in the KKT conditions being equal to zero or not. The resulting systems of equations when solved provide potential candidate solutions to the mathematical program (P). To qualify as a candidate solution, any potential candidate solution will have to meet the feasibility conditions as well as the conditions required by the properties[6] of RZF's nontrivial zeros, as stated in subsection (3) below. Furthermore, candidate solutions have to meet the KKT sufficiency conditions for optimality (minimization of (P)) since these candidates solutions can possibly be maxima or saddle points.

3. Nontrivial zeros conditions

3a. RZF vanishes at $t = t^*$ for some $\sigma = \sigma^*$ that meets the KKT necessary conditions, so that:

$U^* = U(\sigma^*; t^*) = 0$ and $V^* = V(\sigma^*; t^*) = 0$, and based on RZF's property (1), i.e. U^* *and* V^* are differentiable, hence:

$$Z^*_\sigma = U^*U^*_\sigma + V^*V^*_\sigma = 0 \qquad (7)$$

3b. RZF's property (8) requires $\sigma > 0$ (8)

3c. RZF's property (4) requires

Either: Nontrivial zeros on the critical line: $g(\sigma) = \sigma - 1/2 = 0$ (9a)

Or: Pairs of nontrivial zeros off of the critical: $g(\sigma) = \sigma - ½ < 0$ (9b)

Proof

RZF's Property (8) requires $\sigma^* > 0$, so that KKT condition (1) reduces to:

$$Z^*_\sigma + \mu^* = 0 \qquad (1b)$$

Hence, from (7) and (1b) we get $\mu^* = 0$ as a necessary value for μ that meets KKT conditions above.

Under case (9a), where $g(\sigma^*) = 0$, we get $\sigma^* = ½$. Based on RZF's property (4), this result rules out case (9b), i.e. $g(\sigma) < 0$ is not possible, and the value 1/2 is the only possible candidate solution that meets the KKT conditions as well as the properties of RZF's nontrivial zeros .

Hence there exists only one vector $v^* = (\sigma^*=1/2,\ \mu^*= 0)$ which meets KKT's necessary conditions as well as RZF's nontrivial zeros properties, thus v^* is a solution candidate. It remains to prove that this solution is a minimum for problem (P) and not a maximum or a saddle point. This entails proving that the KKT conditions are also sufficient.

The traditional sufficiency condition [5] requires that the Hessian of the Lagrangian $L_{\sigma\sigma} (\sigma^*, \mu^*; t^*)$ be positive definite for all directions $u \neq 0$ that are defined by $ug_\sigma (\sigma^*) \geq 0$, that is for all directions $u \geq 0$, since $g_\sigma (\sigma) = 1$. Hence, since u needs to be nonzero, for any $u > 0$, the sufficiency condition for problem (P) is: $u^T L_{\sigma\sigma} (\sigma^*, \mu^*) u > 0$. In our case, the directions u

[6] *Properties of RZF and those of its nontrivial zeros are listed* in *the appendix*

are univariate since there is only one variable in problem (P), namely σ, hence the sufficiency condition is: $u^2 L^*_{\sigma\sigma} (\sigma^*, \mu^*) > 0$

From $L_\sigma = Z_\sigma + \mu = Z_\sigma = 2(UU_\sigma + VV_\sigma) + \mu$, we get

$$L_{\sigma\sigma} = 2(U_\sigma^2 + UU_{\sigma\sigma} + V_\sigma^2 + VV_{\sigma\sigma}). \quad (10)$$

From condition (3a) we have: $U^* = V^* = 0$, and based on RZF's property (1), $U^*_{\sigma\sigma}$ and $V^*_{\sigma\sigma}$ exist, thus are finite, the Hessian is then $L^*_{\sigma\sigma} = 2u^2 (U^{*2}_\sigma + V^{*2}_\sigma) > 0$, i.e. the Hessian of $L_{\sigma\sigma} (\sigma^*, \mu^*; t^*)$ is positive definite. This proves that the KKT conditions are also sufficient for optimality at $\sigma^* = ½$.

Therefore, the KKT conditions are necessary and sufficient for the minimum of problem (P) to be at $\sigma^* = ½$, for any $t = t^*$ where RZF vanishes. This shows that the nontrivial zeros are all on the critical line as postulated by the Riemann Hypothesis, which is then analytically proven true by our optimization approach.

As a computational validation, we implemented the proposed optimization approach as stated in problem (P) for the first one hundred nontrivial zeros. Knowing the heights t^* where RZF vanishes, the search for the location of nontrivial zeros on the line $t = t^*$ can be done using a simple one-variable grid search over σ in $(0, 1/2)$. The results validate our analytical proof for the location of nontrivial zeros at $\sigma^* = ½$. Also, knowing that the nontrivial zeros are located on the critical line greatly simplifies the search for their location on this line. Indeed, this can be done by minimizing RZF's squared norm over the variable t for $\sigma = ½$ using a simple one-variable grid search. Since this task entails only the computation of RZF's values at a finite number of t values, the optimization approach should prove much faster and more efficient than currently available methods used in computing RZF's nontrivial zeros on the critical line.

A noteworthy observation is that in the proposed approach to identifing the location of RZF's nontrivial zeros, the analysis did not require the use of a closed form expression of RZF and of its derivatives, but used instead a set of RZF's properities which were sufficient to show that RZF's nontrivial zeros are located on the critical line. Hence, the same approach is valid for any complex-valued non-closed form function that has the same properties as RZF[7]. As an example, the Riemann $\xi(s)$ function also has its zeros located on the critical line [6].

Conclusion

Optimization models provided efficient tools for proving several scientific laws and theories. Based on the success of this approach, we modeled the search for the location of the nontrivial zeros of the Riemann Zeta function in an optimization framework.The properties of RZF and those of its nontrivial zeros enabled the formulation of the search for their location as a constrained optimization problem using the simple objective function of minimizing the squared norm of RZF at some heitght $t = t^*$ where it vanishes, under the constraint that nontrivial zeros are located on the left half of the critical strip. The Karush-Kuhn-Tucker

[7] *See the appendix for a list of RZF's properties*

necessary and sufficient optimality conditions of the resulting constrained nonlinear programming problem proved that the nontrivial zeros of RZF are located on the critical line, thus confirming the conjecture stated in the Riemann Hypothesis. This result is also valid for complex-valued functions that have the same properties as RZF.

Appendix: Some relevant Properties of RZF and its nontrivial zeros

The most important and relevant properties of RZF [7] are listed below:

1. Since RZF is analytic in the complex plane except for a pole at $\sigma=1$, its real and its imaginary parts, $U(s)$ and $V(s)$ respectively, are twice differentiable in the critical strip hence:

   $U^*U^*_\sigma + V^*V^*_\sigma = 0$ (1a); and $U^*U^*_{\sigma\sigma} + V^*V^*_{\sigma\sigma} = 0$ (1b)

2. RZF has an infinite number of nontrivial zeros
3. A huge number of nontrivial zeros proved to be located on the critical line
4. As a consequence of the functional equation, nontrivial zeros either occur on the critical line or in pairs off of the critical line symmetrically about it.
5. Nontrivial zeros are located on the critical strip at different heights $t = t^*$
6. Nontrivial zeros are symmetric about the real line $t = 0$, and about the critical line
7. As per (5), if σ^* is a location of a nontrivial zero at $t = t^*$, then $(1- \sigma^*)$ is also a location of a nontrivial zero at $t = t^*$
8. RZF has no zeros on the line $\sigma = 1$. Thus, by symmetry about the critical line, RZF has no zero on the line $\sigma = 0$, hence for nontrivial zeros: $\sigma > 0$
9. $U_\sigma (\sigma = ½) \neq 0$ *and* $V_\sigma (\sigma = ½) \neq 0$
10. Property (3) and (6) limit the search for nontrivial zeros to the left half of the critical strip. This leads to the constraint: $\sigma \leq ½$

References

1. Riemann Hypothesis, *https://en.wikipedia.org/wiki/Riemann_hypothesis.*
2. Bordley, R. (1983). *The Central Principle of Science. Optimization, Behavioral Science, 28(1) (p. 53).*
3. Luptácik, M. (2010). Mathematical Optimization and Economic Analysis. *Springer Optimization and Its Application (Book 36), (Ch. 2, pp.25 – 36).*
4. Luptácik, M. (2010). Mathematical Optimization and Economic Analysis. *Springer Optimization and Its Application (Book 36), (Ch. 2, p.23).*
5. Neos Guide/Server (2019). Nonlinear Programming. https://neos-guide.org/content/nonlinear programming.
6. Chandrasekhar K. (1953). Lectures on the Riemann Zeta–Function. https://julianoliver.com/share/free-science-books/tifr01, *(p. 106)*.
7. Chandrasekhar K. (1953). Lectures on the Riemann Zeta–Function. https://julianoliver.com/share/free-science-books/tifr01, *(pp. 105-109)*.