

Linear Algebraic Number Theory, Part I: Foundations

Joram Soch

BCCN Berlin, Germany

joram.soch@bccn-berlin.de

Abstract

We introduce a new framework called linear algebraic number theory (LANT) that reformulates the number-theoretic problem as a regression model and solves it using matrix algebra. This framework restricts all computations to log space, therefore replaces multiplication with addition and allows to capture variation in the natural numbers from variation in the prime numbers. This automatically puts prime numbers to their designated place of atomic particles of natural numbers and enables fruitful new formulations of number-theoretic functions. We outline the theory, derive some basic results, make connections to standard number theory and give an outlook regarding the Riemann hypothesis, number theory's long-standing enigma.

Contents

1	Introduction	1
2	Definitions	2
3	Modelling the Natural Numbers	8
4	Inverting the Factorization Matrix	12
5	Expressing Number-Theoretic Functions	14
5.1	IsPrime	14
5.2	PrimeCount	14
5.3	Chebyshev functions	14
5.4	von Mangoldt function	15
5.5	Riemann ζ function	15
6	The Riemann Hypothesis	16
7	References	17

1 Introduction

Let n be a positive natural number. The *fundamental theorem of arithmetic* states that there is a unique factorization by which n can be written as a product of prime powers:

$$n = \prod_{i=1}^k p_i^{n_i} . \quad (1)$$

In this product, k is the number of primes that divide n , $p_1 < \dots < p_k$ are prime numbers and n_1, \dots, n_k are positive integers. This is called the *canonical representation* or *standard form* of n . For example,

$$360 = 2^3 \times 3^2 \times 5^1 . \quad (2)$$

We can take the natural logarithm of equation (1) and obtain

$$\ln n = \sum_{i=1}^k n_i \ln p_i . \quad (3)$$

Applied to the example, this gives

$$\ln 360 = 3 \ln 2 + 2 \ln 3 + 1 \ln 5 . \quad (4)$$

We can write equation (3) as a vector product and obtain

$$\ln n = [n_1 \quad \dots \quad n_k] \begin{bmatrix} \ln p_1 \\ \vdots \\ \ln p_k \end{bmatrix} . \quad (5)$$

Applied to the example, this gives

$$\ln 360 = [3 \quad 2 \quad 1] \begin{bmatrix} \ln 2 \\ \ln 3 \\ \ln 5 \end{bmatrix} . \quad (6)$$

Note that, as $p^0 = 1$ for any $p \in \mathbb{R}, p \neq 0$, one might insert prime powers with exponent zero in equation (1) or log primes with factor zero in equation (3) without changing the value of n . This means that the second vector in equation (5) will be the same for all n , namely a column vector of all log primes, and just the first vector has to be adapted in order to achieve the correct combination of primes.

The basic idea of this paper is to make use of this insight and rewrite the prime factorization of natural numbers (or, log-prime summation of log integers) as a linear equation system which has the logarithmized natural numbers on its left-hand side and a matrix product of a factorization matrix and the log primes on its right-hand side. This is referred to as *linear algebraic number theory* (LANT).

2 Definitions

An example for such a linear equation system is

$$\begin{aligned}
 \ln 1 &= 0 \ln 2 + 0 \ln 3 + 0 \ln 5 \\
 \ln 2 &= 1 \ln 2 + 0 \ln 3 + 0 \ln 5 \\
 \ln 3 &= 0 \ln 2 + 1 \ln 3 + 0 \ln 5 \\
 \ln 4 &= 2 \ln 2 + 0 \ln 3 + 0 \ln 5 \\
 \ln 5 &= 0 \ln 2 + 0 \ln 3 + 1 \ln 5 \\
 \ln 6 &= 1 \ln 2 + 1 \ln 3 + 0 \ln 5
 \end{aligned} \tag{7}$$

which, in matrix algebra notation, can be written as

$$\begin{bmatrix} \ln 1 \\ \ln 2 \\ \ln 3 \\ \ln 4 \\ \ln 5 \\ \ln 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \ln 2 \\ \ln 3 \\ \ln 5 \end{bmatrix}. \tag{8}$$

In order to formulate this for the general case, we will introduce some definitions.

Definition 1: (*element-wise logarithm*) Whenever the natural logarithm is applied to a vector $v \in \mathbb{R}^n$, it is calculated element-wise:

$$\ln v = \ln \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \ln v_1 \\ \vdots \\ \ln v_n \end{bmatrix}. \tag{9}$$

Definition 2: (*natural number vector*) Let n be a positive natural number. Then, the $n \times 1$ vector z_n is defined as

$$z_n = \begin{bmatrix} 1 \\ 2 \\ \vdots \\ n \end{bmatrix}. \tag{10}$$

Definition 3: (*prime number vector*) Let n be a positive natural number. Then, the $\pi(n) \times 1$ vector p_n is defined as

$$p_n = \begin{bmatrix} 2 \\ 3 \\ 5 \\ \vdots \\ p \end{bmatrix} \tag{11}$$

where p is the largest $x \in \mathbb{P}$ for which $x \leq n$, \mathbb{P} is the set of prime numbers and $\pi(n)$ is the number of primes less than or equal to n .

Obviously, as implied by the fundamental theorem of arithmetic (1) and instantiated by the above example (8), z_n and p_n can be related to each other in log space by a matrix of coefficients. We will call this the *factorization matrix*.

Definition 4: (*prime factorization matrix*) Let n be a positive natural number. Then, the prime factorization matrix is the $n \times \pi(n)$ matrix F_n for which

$$\ln z_n = F_n \ln p_n . \tag{12}$$

We know that F_n exists for a given $n > 1$, because every natural number greater 1 is either prime or can be factorized into primes smaller than itself, and equation (12) is nothing but a restatement of that fact. We also know that F_n is unique as there is only one prime factorization for every $n > 1$ which can be proven with recourse to Euclid's lemma (Euclid, VII, 30). An example for $n = 20$ is given in Figure 1.

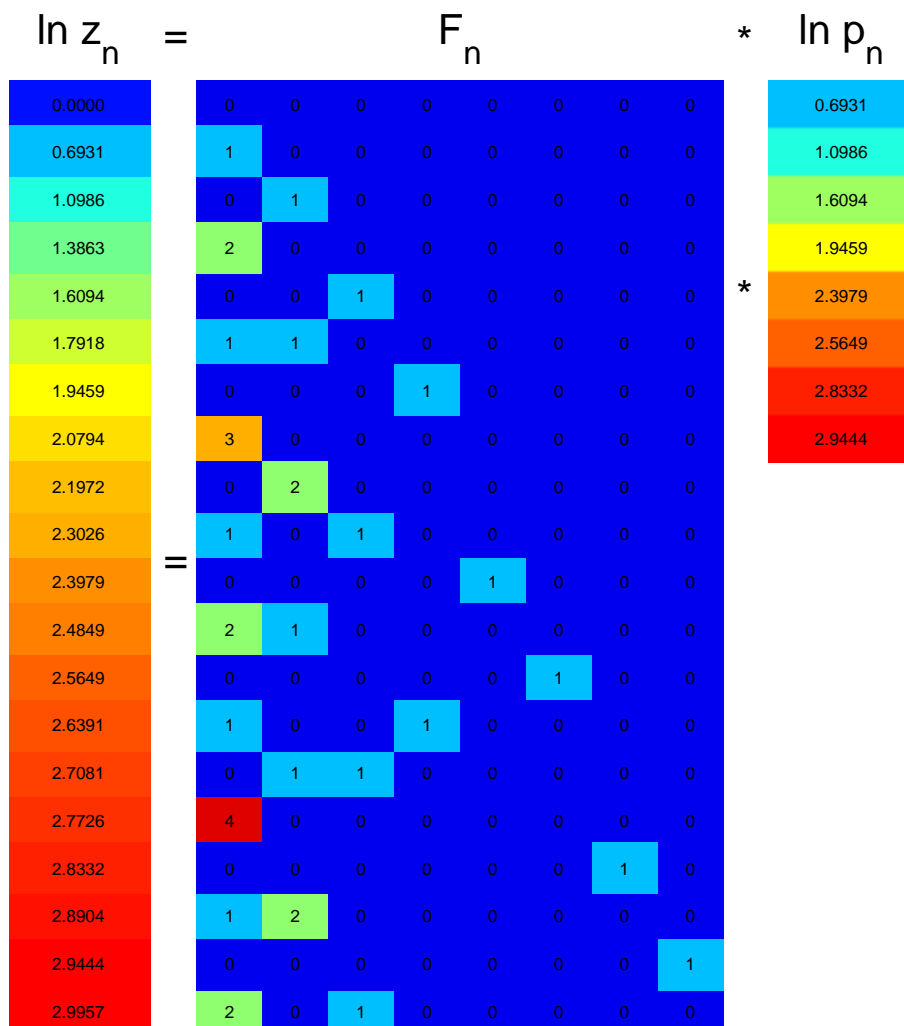


Figure 1. Log integers $\ln z_n$, factorization matrix F_n and log primes $\ln p_n$ for $n = 20$. This figure illustrates the log-space analogue of integer factorization in which log integers are represented as sums of log primes, weighted by the factorization matrix.

Suppose we didn't know the primes up to a certain number n . A natural consequence would be that we try to derive, solve for, infer on or estimate them. To this end, we introduce the concept of *candidate primes*.

Definition 5: (*candidate prime vector*) Let n be a positive natural number. A candidate prime vector q is an $m \times 1$ vector with $m \leq n$ which only contains pairwise different natural numbers smaller than or equal to n .

For example, the following would be possible candidate primes for $n = 10$:

$$q_1 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \quad q_2 = \begin{bmatrix} 2 \\ 3 \\ 5 \\ 7 \end{bmatrix} = p_{10}, \quad q_3 = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \\ 7 \\ 8 \end{bmatrix}, \quad q_4 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{bmatrix} = z_{10}. \quad (13)$$

It would now be tempting to take a certain vector q and somehow calculate its associated matrix F_n in order to factorize potentially large numbers. However, this is neither possible nor necessary. It is not possible as this equation system would contain more unknowns than equations. It is not necessary as every set of candidate primes already implies a factorization matrix which we call a *candidate factorization*.

Definition 6: (*candidate factorization matrix*) Let q be a candidate prime vector. The candidate factorization matrix $F_n(q)$ is the $n \times m$ matrix that would be the prime factorization matrix, if q were the true primes.

For example, as 2 is a prime number, every second number is factorized by 2^1 , every fourth number is factorized by 2^2 , every eighth number is factorized by 2^3 and so on. Consequently, the first column of F_n has a 1 in every second row, a 2 in every fourth row, a 3 in every eighth row (see Figure 1). Similarly, if 4 were prime (which it is not), the corresponding column of F_n would have a 1 in every fourth row, a 2 in every sixteenth row, a 3 in every sixty-fourth row etc.

To continue with the example from above, the candidate factorizations for the candidate primes in equation (13) would be:

$$F_{10}(q_1) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 2 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ 3 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad F_{10}(q_2) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad F_{10}(q_3) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 3 & 0 & 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad (14)$$

$$F_{10}(q_4) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (15)$$

To express general candidate factorizations, we will introduce some more definitions.

Definition 7: (*basic vectors*) The zero vector and the ones vector:

$$0_n = \left. \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \right\} \text{ n zeros}, \quad 1_n = \left. \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \right\} \text{ n ones}. \quad (16)$$

Definition 8: (*elementary vectors*) The i -th elementary vector in n -dimensional vector space is an n -dimensional zero vector with a one in its i -th entry:

$$e_{i|n} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \leftarrow i\text{-th position} \\ \\ \leftarrow n\text{-th position} \end{array}. \quad (17)$$

Definition 9: (*periodic elementary vectors*) The i -th periodic elementary vector in n -dimensional vector space is an n -dimensional zero vector with a one in its k -th entries where $k = i, 2i, \dots, \lfloor n/i \rfloor i$:

$$e_{\bar{i}|n} = \sum_{k=1}^{\lfloor n/i \rfloor} e_{k \cdot i | n} = \begin{bmatrix} 1_{\lfloor n/i \rfloor} \otimes e_{i|i} \\ 0_{\text{mod}(n,i)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \leftarrow i\text{-th position} \\ \\ \leftarrow 2i\text{-th position} \\ \\ \leftarrow n\text{-th position} \end{array}. \quad (18)$$

With these definitions, we are now able to express arbitrary factorization matrices:

Definition 10: (*factorization vector*) The i -th factorization vector in n -dimensional vector space describes how often i would occur as a factor in the prime factorization of the numbers z_n , if i were prime. It is given by

$$f_{i|n} = \sum_{j=1}^{\lfloor \log_i n \rfloor} e_{i^j|n} . \quad (19)$$

By definition, we set

$$f_{1|n} = 1_n . \quad (20)$$

Theorem 1: Let q be a candidate prime vector. Then, the corresponding candidate factorization matrix is given by

$$F_n(q) = [f_{q_1|n} \cdots f_{q_m|n}] . \quad (21)$$

Proof 1: This follows from Def. 6 and 10. The sum over periodic elementary vectors in equation (19) ensures that the j -th power of each candidate prime q_i repeats every q_i^j -th entry, because q_i^j would be part of these factorizations, if q_i was prime. ■

3 Modelling the Natural Numbers

The factorization matrix is at the heart of LANT and appears in its fundamental theorem:

$$\ln z_n = F_n \ln p_n . \quad (26)$$

We will now again assume that p_n is unknown, so that we have to solve for it. This can be nicely connected to linear models and statistical modelling as, when searching the optimal solution for $\ln p_n$, we are seeking the best way to capture the values of the natural numbers, just like we are seeking the best way to capture the variance in measured data when applying linear models to empirical phenomena.

The univariate linear regression model is given by

$$y = X\beta + \varepsilon . \quad (27)$$

In this equation, certain data ($n \times 1$ vector y) are modelled as a linear combination of independent variables ($n \times p$ matrix X), weighted by some coefficients ($p \times 1$ vector β), plus some residuals that cannot be explained ($n \times 1$ vector ε) where y is called the *signal*, X is called the *design matrix*, β are called *regression coefficients* and ε is called *noise*. We observe the following parallels between (26) and (27):

- The log integers $\ln z_n$ are the signal y that we want to explain.
- The factorization matrix F_n is the design matrix X that we use to explain.
- The log primes $\ln p_n$ are the regression coefficients β that we want to estimate.
- If we use the prime factorization matrix F_n as the design matrix, there are no residuals ε as the natural numbers are completely explained by the prime numbers (see Figure 1). However, if we use a candidate factorization matrix lacking some primes, we will fail to resolve the complete variation in the natural numbers, so that there will be errors ε (see Figure 2).

We can therefore write down the statistical version of equation (26):

$$\ln z_n = F_n(q) \ln q + \varepsilon_n . \quad (28)$$

With the concepts of candidate primes and candidate factorization, we have a simple method of constructing the design matrix for our linear regression (28). The next step is therefore to estimate the model, i.e. to find some parameters, given the data and the design:

$$\hat{\beta} = f(y, X) . \quad (29)$$

Naturally, when performing linear regression, one wants to keep the residuals ε as small as possible in order to achieve “the best possible fit” of the model to the data. A common framework for assigning parameter values following this rationale is *ordinary least squares* (OLS).

Definition 13: (*ordinary least squares*) Let z_n be the natural numbers up to n . Further, consider candidate primes q and the candidate factorization $X = F_n(q)$. Then

- 1) $\ln z_n = F_n(q) \ln q + \varepsilon_n$ is called a “linear factorization model” of z_n ;
- 2) $\ln \hat{q} = (X^T X)^{-1} X^T (\ln z_n)$ is called the “log-prime estimator” (LPE);
- 3) $\ln \hat{q}$ are also referred to as the “estimated log primes”;
- 4) $\ln \hat{z}_n = F_n(q) \ln \hat{q}$ are called the “predicted log integers”.

Theorem 2: The LPE minimizes the residual sum of squares.

Proof 2: The residual sum of squares for (27) is given by

$$\text{RSS}(\beta) = \sum_{i=1}^n \varepsilon_i^2 = \varepsilon^T \varepsilon = (y - X\beta)^T (y - X\beta) \quad (30)$$

which can be expanded to

$$\text{RSS}(\beta) = y^T y - y^T X\beta - \beta^T X^T y + \beta^T X^T X\beta \quad (31)$$

and differentiated to

$$\text{RSS}'(\beta) = 2X^T X\beta - 2X^T y. \quad (32)$$

Setting this derivative to zero yields

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (33)$$

which conforms to the estimated log primes in Def. 13.2. ■

With the OLS estimator at hand, we can now consider different cases of candidate primes:

- Case I: The candidate primes are a real subset of the prime numbers: $q \subset p_n$.
- Case II: The candidate primes equal the prime numbers: $q = p_n$.
- Case III: The candidate primes are a real superset of the prime numbers: $q \supset p_n$.
- Case IV: The candidate primes equal the natural numbers: $q = z_n$.

Note that these cases generalize the examples from equation (13). Figure 2 shows one example for each case and compares (i) the log natural numbers $\ln z_n$ to the predicted log integers $\ln \hat{z}_n$ as well as (ii) the log candidate primes $\ln q$ to the estimated log primes $\ln \hat{q}$, as given in Def. 13. The candidate primes used in the figure are:

- Case I: $q_1 = [3, 5, 11, 17]^T$.
- Case II: $q_2 = [2, 3, 5, 7, 11, 13, 17, 19]^T = p_{20}$.
- Case III: $q_3 = [2, 3, 4, 5, 7, 8, 11, 12, 13, 15, 17, 19]^T$.
- Case IV: $q_4 = [1, 2, 3, \dots, 18, 19, 20]^T = z_{20}$.

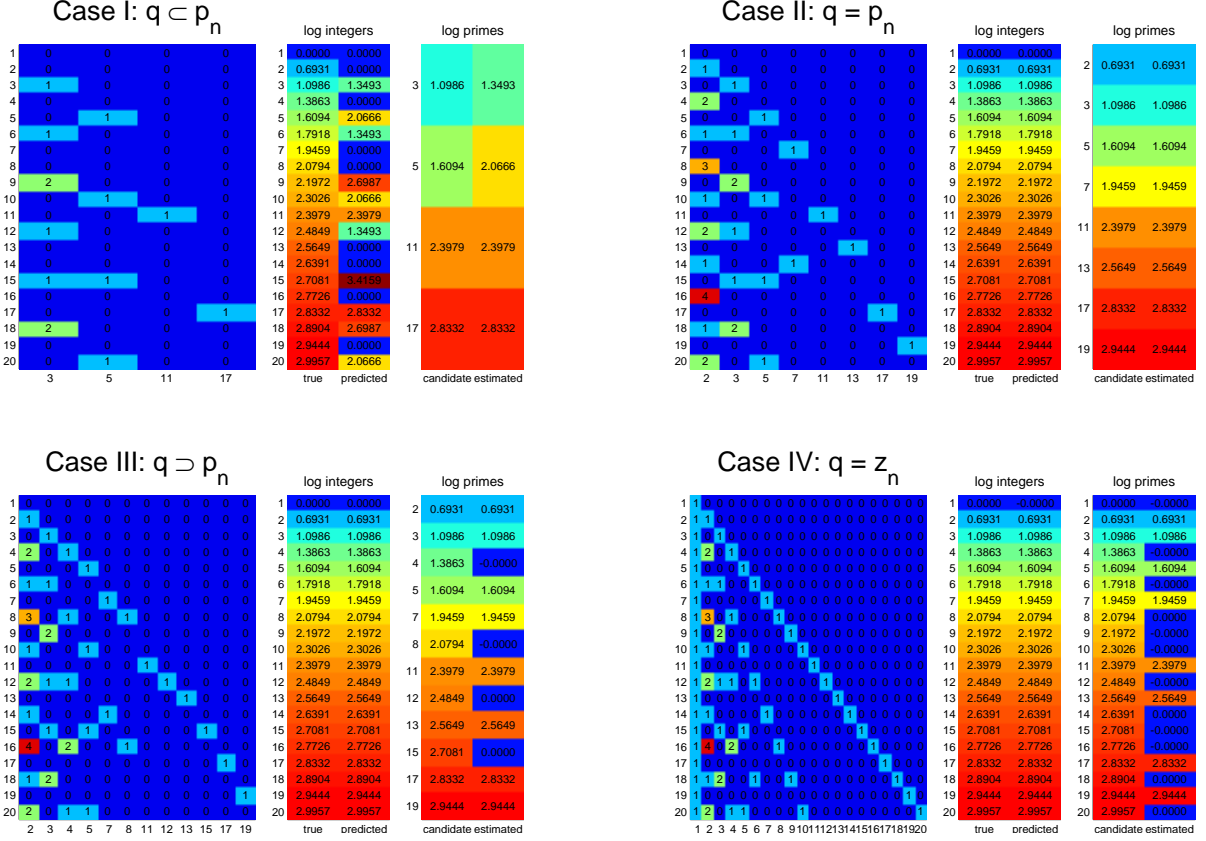


Figure 2. Four different cases of candidate primes for $n = 20$. Candidate prime vectors are given in the text. Case II is also used in Figure 1 and Case IV is also used in Figure 3. Each panel consists of the candidate factorization (left), comparison of log natural numbers $\ln z_n$ vs. predicted log integers $\ln \hat{z}_n$ (middle) and comparison of log candidate primes $\ln q$ vs. estimated log primes $\ln \hat{q}$ (right). All in all, we make the following observations: (i) As soon as all primes smaller than or equal to n are included in q , the log integers are predicted perfectly with maximal accuracy (upper right and lower panels) in which case we call q “complete”. If some prime numbers are missing, not all variation in the natural numbers can be captured (upper left panel). (ii) Only if $q = p_n$, the candidate primes are identical to the estimated primes (upper right panel) in which case we call q “valid”. If some primes are missing or non-primes are present, there is disagreement (upper left and lower panels). (iii) If q contains all primes smaller than or equal to n , non-primes are automatically “switched off” by the LPE and receive a weight of zero whereas primes receive their logarithm as weight (lower panels), consistent with the fundamental theorem of arithmetic (1) and its logarithmized version (3). If some elements of p_n are missing in q , estimation tends to be unreliable (upper left panel), consistent with the view of the primes as the atomic particles of the natural numbers. (iv) In summary, one can say that the prime numbers are the sparsest set using which one can fully decompose the natural numbers (upper right panel). Equivalently, one could say that the primes are those numbers from the set of all possible candidate primes that minimize the *prediction error* $(\ln z_n - \ln \hat{z}_n)^T (\ln z_n - \ln \hat{z}_n)$ and the *estimation error* $(\ln q - \ln \hat{q})^T (\ln q - \ln \hat{q})$. This refines prime number identification as a model comparison problem in which the least complex from the most accurate models is selected as the optimal solution.

Based on these observations, we set up consistency conditions for candidate primes and formulate a theorem about the behavior of the LPE for different candidate primes.

Definition 14: (*consistency conditions*) Let q be candidate primes and $\ln \hat{q}$ the LPE. Then, we call q

- 1) “valid”, if $\ln q = \ln \hat{q}$;
- 2) “complete”, if $\ln z_n = \ln \hat{z}_n$;
- 3) “consistent”, if it is valid and complete.

Theorem 3: Let q be candidate primes and $X = F_n(q)$ the corresponding candidate factorization. Then, $\ln \hat{q} = (X^T X)^{-1} X^T (\ln z_n)$ and:

- 1) If q is consistent, then $q = p_n$ and vice versa.
- 2) If $q = p_n$, then $\ln \hat{q} = \ln p_n$.
- 3) If $q \supset p_n$, then

$$(\ln \hat{q})_j = \begin{cases} \ln q_j & , \text{ if } q_j \in \mathbb{P} \\ 0 & , \text{ if } q_j \notin \mathbb{P} \end{cases}, \quad j = 1, \dots, m. \quad (34)$$

- 4) If $q = z_n$, then

$$(\ln \hat{q})_i = \begin{cases} \ln i & , \text{ if } i \in \mathbb{P} \\ 0 & , \text{ if } i \notin \mathbb{P} \end{cases}, \quad i = 1, \dots, n. \quad (35)$$

Proof 3: We prove this theorem step by step.

1) If q is consistent, it follows from Def. 14.1, 14.2 and 13.4 that $\ln z_n = F_n(q) \ln q$. According to Def. 4, there is only one solution for q and this is $q = p_n$. Conversely, if $q = p_n$, then $F_n(q) = F_n$ by Def. 6. We also know that $\ln z_n = F_n \ln p_n$ from Def. 4 which implies that $p_n = q$ is consistent according to Def. 14.3. \square

2) If $q = p_n$, then $F_n(q) = F_n$ by Def. 6. For this case, Def. 4 gives a solution for which $\varepsilon_n = 0_n$, namely $\ln q = \ln p_n$. If there is a solution for which $\text{RSS}(\ln q) = 0$, the LPE must select this solution by Th. 2. Therefore, $\ln \hat{q} = \ln p_n$. \square

3) If $q \supset p_n$, q contains primes and non-primes. For this case, we can construct a solution for which $\varepsilon_n = 0_n$, namely the solution given by (34). By Def. 5 and Th. 1, columns of $F_n(q)$ are linearly independent. Therefore, this is the only solution for which $\text{RSS}(\ln q) = 0$. The rest follows the proof of 2). \square

4) This is a special case of 3). \square

This completes the proof. \blacksquare

4 Inverting the Factorization Matrix

In this section, we want to develop something like a *collective primality test* for the set of all natural numbers up to n with the help of the following theorem:

Theorem 4: If $q = z_n$, then $F_n(q)$ is a quadratic $n \times n$ matrix and

- 1) $F_n(q)$ is invertible;
- 2) $\ln \hat{q} = [F_n(q)]^{-1} (\ln z_n)$.

Proof 4: We prove this theorem step by step.

1) From Def. 6 and Th. 1, it follows that $F_n(z_n)$ is a lower triangular matrix. The determinant of a triangular matrix equals the product of its diagonal entries. Since all diagonal elements of $F_n(z_n)$ are 1, $\det [F_n(z_n)] = 1 \neq 0$ and $F_n(z_n)$ is invertible. \square

2) The theory of linear equation systems states that an LES that can be represented as $Ax = b$ with the $n \times n$ invertible matrix A has exactly one solution given by $\hat{x} = A^{-1}b$. Translated to our example, this implies that $\ln \hat{q} = [F_n(z_n)]^{-1} (\ln z_n)$. \square

This completes the proof. \blacksquare

As we now know (i) that the LPE, quite comfortably, “switches off” non-prime entries in a candidate prime vector (see Theorem 3.4) and (ii) that the LPE reduces to a simpler form when the candidate primes equal the natural numbers (see Theorem 4.2), the problem really reduces to finding the inverse of $F_n(z_n)$. We put forward the following solution:

Theorem 5: If $q = z_n$, then $F_n(q)$ is a quadratic $n \times n$ matrix and

$$[F_n(q)]^{-1} = \prod_{i=1}^n \left(M_{n+1-i}(2) - \left(\sum_{j=1}^{\lfloor \log_{(n+1-i)} n \rfloor} e_{(n+1-i)j|n} \right) e_{n+1-i|n}^T \right). \quad (36)$$

Proof 5: We will use Gauss-Jordan elimination to invert the factorization matrix. This means, we will transform $F_n(z_n)$ into the identity matrix I_n by left-multiplication with elementary matrices and get $[F_n(z_n)]^{-1}$ as the product of these matrices.

Let $F = F_n(z_n)$. According to Th. 1 and Def. 10, the i -th column of F is

$$f_{i|n} = \sum_{j=1}^{\lfloor \log_i n \rfloor} e_{i\bar{j}|n}. \quad (37)$$

In order to remove $f_{i|n}$ from F , i.e. replace it by $e_{i|n}$ to reach I_n , we have to left-multiply F with a matrix containing $-f_{i|n}$. However, (i) $-f_{i|n}$ has to be placed into the i -th column and (ii) it may not remove the diagonal 1 from

F . This is achieved by (i) right-multiplying $-f_{i|n}$ with the transposed $e_{i|n}$ and (ii) subtracting it from the I_n that has a 2 in the i -th column. In this way, we obtain the extended elementary matrix E_i :

$$E_i = M_i(2) - f_{i|n} e_{i|n}^T. \quad (38)$$

If F is successively left-multiplied with E_i , $i = 1, \dots, n$, it will become I_n . Since we have to reverse the order in the product to account for successive left-multiplication, we obtain:

$$F^{-1} = \prod_{i=1}^n E_{n+1-i} = \prod_{i=1}^n (M_{n+1-i}(2) - f_{n+1-i|n} e_{n+1-i|n}^T). \quad (39)$$

With application of (37), we have

$$F^{-1} = \prod_{i=1}^n \left(M_{n+1-i}(2) - \left(\sum_{j=1}^{\lfloor \log_{(n+1-i)} n \rfloor} e_{(n+1-i)^j|n} \right) e_{n+1-i|n}^T \right) \quad (40)$$

which conforms to equation (36). ■

Figure 3 displays an example for inversion of a quadratic candidate factorization matrix.

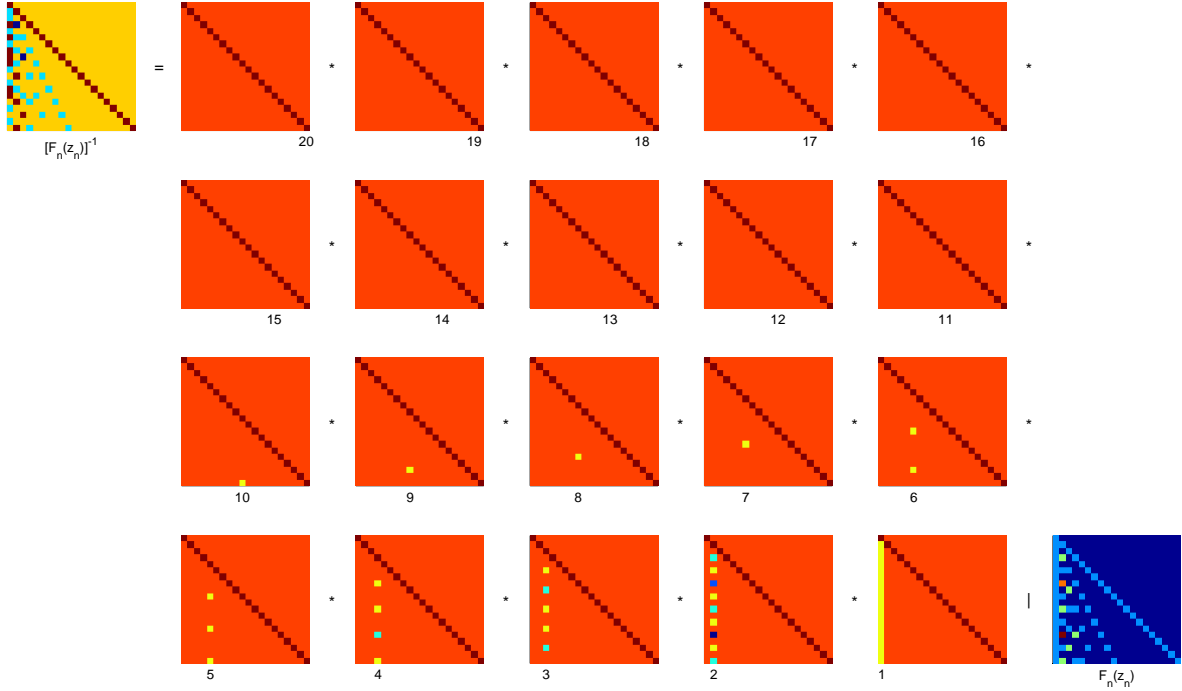


Figure 3. Example for inverting the factorization matrix with $n = 20$. The candidate factorization $F_n(z_n)$ is shown in the lower right, the extended elementary matrices for left-multiplication E_n, \dots, E_1 are shown in the middle and the resulting matrix inverse $[F_n(z_n)]^{-1}$ is shown in the upper left. Note that the matrices have a different color scale.

5 Expressing Number-Theoretic Functions

In this section, we want to employ the results derived so far, especially Theorem 3, to express certain functions that are important in the field of number theory.

5.1 IsPrime

The *IsPrime function* $\text{ip}(x)$ is defined as (OEIS, A010051)

$$\text{ip}(x) = \begin{cases} 1 & , \text{ if } x \in \mathbb{P} \\ 0 & , \text{ if } x \notin \mathbb{P} \end{cases} . \quad (41)$$

Using LANT terminology, $\text{ip}(x)$ can be expressed amazingly simple as

$$\text{ip}(i) = \frac{(\ln \hat{q})_i}{\ln i} \quad \text{where} \quad \ln \hat{q} = [F_n(z_n)]^{-1} (\ln z_n) \quad \text{with} \quad n \geq i \quad (42)$$

which is a trivial consequence of Theorem 3.4. Note that $\text{ip}(0)$ is not defined and that $\text{ip}(1)$ is an indeterminate form, consistent with 1 being considered neither prime nor composite. Further, in contrast to $\text{ip}(x)$, $\text{ip}(i)$ is only defined for positive natural numbers.

5.2 PrimeCount

The *PrimeCount function* $\pi(x)$ is defined as (OEIS, A000720)

$$\pi(x) = |\{n \in \mathbb{P} \mid n \leq x\}| . \quad (43)$$

In the LANT framework, $\pi(x)$ can be expressed similarly simple as

$$\pi(n) = \sum_{i=2}^n \text{ip}(i) = \sum_{i=2}^n \frac{(\ln \hat{q})_i}{\ln i} \quad (44)$$

which follows from equation (42). Note that this sum starts at $i = 2$, because $\text{ip}(1)$ is an indeterminate form. Again, in contrast to $\pi(x)$ which is defined for any number $x \in \mathbb{R}$ (Platt, 2013), $\pi(n)$ is only defined for positive natural numbers.

5.3 Chebyshev functions

The *first Chebyshev function* is given by (Dusart, 2010)

$$\vartheta(x) = \sum_{p \leq x} \ln p \quad (45)$$

and the *second Chebyshev function* is given by (Dusart, 2010)

$$\psi(x) = \sum_{p^k \leq x} \ln p \quad (46)$$

where the sums are extending over all prime numbers $p \in \mathbb{P}$ satisfying $p \leq x$ or $p^k \leq x$.

Again following Theorem 3.4 and based on equation (42), we have

$$\vartheta(n) = \sum_{i=1}^n (\ln \hat{q})_i = 1_n^T (\ln \hat{q}) \quad (47)$$

as a simple expression for the first Chebyshev function. The second Chebyshev function cannot be easily represented using LANT quantities, but is related to the first one by

$$\psi(x) = \sum_{n=1}^{\infty} \vartheta(x^{1/n}) . \quad (48)$$

5.4 von Mangoldt function

The *von Mangoldt function* is given by (Conrey, 2003)

$$\Lambda(n) = \begin{cases} \ln p & , \text{ if } n = p^k, p \in \mathbb{P}, k \geq 1 \\ 0 & , \text{ otherwise} \end{cases} . \quad (49)$$

It can be related to the Chebyshev functions by (Conrey, 2003)

$$\psi(x) = \sum_{n \leq x} \Lambda(n) = \sum_{n=1}^{\lfloor x \rfloor} \Lambda(n) . \quad (50)$$

Using LANT, we obtain the following reformulation of $\Lambda(n)$

$$\Lambda(i) = \sum_{j=1}^n (\ln \hat{q})_j [\ln i = [F_n(z_n)]_{i,j} (\ln \hat{q})_j] \quad (51)$$

where $[a = b]$ is Iverson bracket notation and $[A]_{i,j}$ refers to the (i, j) -th entry of A .

5.5 Riemann ζ function

The previously mentioned functions $\pi(x)$, $\vartheta(x)$, $\psi(x)$ and $\Lambda(n)$ are closely related to the complex-valued *Riemann ζ function* that is given by (Riemann, 1859)

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \dots \quad (52)$$

which, due to the fundamental theorem of arithmetic, is equivalent to

$$\zeta(s) = \prod_{p \in \mathbb{P}} \frac{1}{1 - p^{-s}} = \frac{1}{1 - 2^{-s}} \cdot \frac{1}{1 - 3^{-s}} \cdot \frac{1}{1 - 5^{-s}} \cdot \dots \quad (53)$$

Note that these equations only hold for $\text{Re}(s) > 1$, but $\zeta(s)$ can be analytically continued to the complete real-positive complex half-plane using a Dirichlet eta series by

$$\zeta(s) = \frac{1}{1 - 2^{1-s}} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s} = \frac{1}{1 - 2^{1-s}} \cdot \left(\frac{1}{1^s} - \frac{1}{2^s} + \frac{1}{3^s} - \dots \right) . \quad (54)$$

6 The Riemann Hypothesis

The Riemann ζ function has trivial zeros at $s = -2, -4, -6, \dots$ and non-trivial zeros which are known to lie in the critical strip $0 < \operatorname{Re}(s) < 1$. The *Riemann hypothesis* (RH) states that all non-trivial zeros are located on the critical line with real part $1/2$:

$$s \in \{z \in \mathbb{C} \mid \zeta(z) = 0 \wedge \operatorname{Re}(z) > 0\} \Rightarrow \operatorname{Re}(s) = \frac{1}{2}. \quad (55)$$

RH remains one of number theory's unsolved problems, as it has neither been proven nor falsified so far. However, RH has a lot of important consequences in number theory and is connected to the prime-counting function $\pi(x)$. In particular, it has been shown that RH is equivalent to the following statement (von Koch, 1901):

$$\pi(x) = \operatorname{Li}(x) + \mathcal{O}(\sqrt{x} \ln x). \quad (56)$$

This means that, for a certain $k > 0$ and $x_0 \in \mathbb{R}$, it holds that

$$|\pi(x) - \operatorname{Li}(x)| \leq k \cdot \sqrt{x} \ln x \quad \text{for all } x \geq x_0. \quad (57)$$

Specifically, it has been shown that under RH (Schoenfeld, 1976)

$$|\pi(x) - \operatorname{Li}(x)| \leq \frac{1}{8\pi} \cdot \sqrt{x} \ln x \quad \text{for all } x \geq 2657. \quad (58)$$

In these formulas, $\operatorname{Li}(x)$ is the logarithmic integral function:

$$\operatorname{Li}(x) = \int_2^x \frac{1}{\ln t} dt. \quad (59)$$

Remember that we have an explicit formula for $\pi(x)$ (44) and note how structurally similar this equation is to $\operatorname{Li}(x)$ (59). Therefore, proving RH through means of LANT might be a promising direction. At first sight, there seem to be two strategies:

- Simplify the left-hand side of (57) by writing $\pi(x)$ as an integral.
- Simplify the left-hand side of (57) by writing $\operatorname{Li}(x)$ as a sum.

Incidentally, we have also observed that the function $\operatorname{ld}(n) = \ln(\det[F_n^T F_n])$ seems to be asymptotically equivalent to $\operatorname{Li}(n)$. This conjecture and other questions will be investigated in future research. A good point to start with might be the further inversion of the factorization matrix (Soch, in prep.).

7 References

- [1] Conrey JB (2003): “The Riemann Hypothesis”. *Notices of the American Mathematical Society*, vol. 50, no. 3, pp. 341-353.
- [2] Dusart P (2010): “Estimates of some functions over primes without R.H.”. *arXiv math*, arXiv:1002.0442v1; URL: <http://arxiv.org/abs/1002.0442v1>.
- [3] Euclid (1996): *Die Elemente. Bücher I-XIII*. Appeared in *Ostwalds Klassiker der exakten Wissenschaften*, vol. 235. Translated from Greek, edited by Clemens Thaer, with a preface from Wolfgang Trageser, reprinted in 1996, 2nd edition.
- [4] Platt DJ (2013): “Computing $\pi(x)$ analytically”. *arXiv math*, arXiv:1203.5712v3; URL: <http://arxiv.org/abs/1203.5712v3>.
- [5] Riemann B (1859): “Ueber die Anzahl der Primzahlen unter einer gegebenen Größe”. *Monatsberichte der Berliner Akademie*, November 1859.
- [6] Schoenfeld L (1976): “Sharper bounds for the Chebyshev Functions $\theta(x)$ and $\psi(x)$. II”. *Mathematics of Computation*, vol. 30, no. 134, pp. 337-360; DOI: 10.2307/2005976.
- [7] Soch J (in prep.): “Linear Algebraic Number Theory, Part II: The Inverse Factorization”, in preparation.
- [8] von Koch H (1901): “Sur la distribution des nombres premiers”. *Acta Mathematica*, vol. 24, iss. 1, pp. 159-182; DOI: 10.1007/BF02403071.
- [9] *The On-Line Encyclopedia of Integer Sequences* (OEIS); URL: <http://oeis.org/>.