

Theory of Regression Lines and Associated Linear Systems and Channels

Emilio Matricciani

Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, Milan, Italy

Email: Emilio.Matricciani@polimi.it

How to cite this paper: Matricciani, E. (2026) Theory of Regression Lines and Associated Linear Systems and Channels. *Open Journal of Statistics*, 16, 401-418. <https://doi.org/10.4236/ojs.2026.165020>

Received: July 30, 2026

Accepted: September 27, 2026

Published: September 30, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

I have summarized the results of a theory regarding linear regression, previously scattered in several articles, with several applications to real-world data. The linear relationship between a dependent variable y and an independent variable x is interpretable as describing the input-output characteristic of a linear system, or a linear channel. The conviction that the theory is applicable to any scientific discipline—provided a linear input-output relationship exists—has led me to summarize the main results in a general form easily applicable to any experimental problem. Defining the regression noise-to-signal power ratio (NSR) R_m and the correlation noise-to-signal power ratio, R_r , the channel noise-to-signal ratio R is the sum of the two. A direct and insightful analysis is achieved by using the NSR, instead of the reciprocal signal-to-noise ratio (SNR), because it makes easier to study which addend determines R and allows a useful graphical representation. The SNR can be considered an indicator of the degree of deterministic nature of a channel, because the higher it is, the greater the “degree” of the channel’s deterministic nature. I have considered various connections of single channels (single, cross, deterministic, parallel, series) that are likely found in applications. In conclusion, the theory and its mathematical results concerning the input-output characteristics of linear channels, described by experimental regression lines, should be useful in any scientific discipline.

Keywords

Regression Line, Noise-to-Signal Ratio, Signal-to-Noise Ratio, Linear Channels, Linear Systems, Cross Channel, Parallel Channel, Series Channel

1. Introduction: Regression Lines Can Model Linear Systems and Channels

Scientific research frequently employs scatter plots in which the relationship be-

tween the dependent variable y and the independent variable x can be modeled linearly—that is, by means of a regression line characterized by a slope m and a regression coefficient r of the sample considered.

This article aims to summarize the results of a theory regarding linear regression, whose relationships can always be interpreted as describing the input-output characteristic of a linear system. In the following discussion, I will refer to “linear channels” instead of “linear systems” because of the terminology I use in the article; however, the reader will appreciate that these are simply two different ways of describing the same functional relationship.

I have developed the theory in a series of articles concerning the mathematical properties of literary texts—both ancient and modern—particularly in relation to translations [1]-[17]. Another important application was reported in Reference [18].

The conviction that the theory is applicable to any scientific discipline—provided a linear input-output relationship exists—has led me to summarize the main results in a general form that the readers can easily apply to their specific problem. In any case, in the articles mentioned, there are numerous experimental applications demonstrating how the theory can be applied to real-world data.

After this introduction, Section 2 defines the single channel and its noise-to-signal power ratio; Section 3 presents and discusses the geometrical representation of the noise-to-signal power ratio; Section 4 defines the cross channel; Section 5 compares a stochastic channel versus its deterministic version; Section 6 studies the parallel connection of channels; Section 7 studies the series connection of channels; Section 8 reports a summary and a conclusion. **Appendix A** lists the mathematical symbols used in the article and their meaning (**Table A1**); **Appendix B** proves a fundamental relationship; **Appendix C** details the calculation of a formula of the main text.

2. Single Channel and Noise-to-Signal Power Ratio

In this section, I recall the general theory on stochastic variables linearly connected [19]. Let y (dependent variable) and x (independent variable) linked by the line:

$$y = m_o x \quad (1)$$

Equation (1) models a *deterministic* relationship through the slope m_o . The intercept $q = 0$; however, if $q \neq 0$, the theory can be fully applied by defining a new dependent variable:

$$\tilde{y} = y - q \quad (2)$$

In the real world, the relationship between y and x is never deterministic, *i.e.*, given by Equation (1), but stochastic (random). Equation (1) models, in fact, two variables perfectly correlated, with correlation coefficient $|r| = 1$ and slope m_o . As these conditions are never found in any application, Equation (1) transforms into:

$$y = mx + n. \quad (3)$$

In Equation (3), $m \neq m_o$ is a multiplicative distortion (bias); n is an additive stochastic variable, with zero mean value, which renders the linear channel partially stochastic, namely “noisy”, using a communication system terminology [20]. In other words, Equation (3) models a noisy channel with multiplicative bias.

Figure 1 shows the flow chart describing Equation (3) with a system/channel representation. The black box indicated with m represents a distorted deterministic channel; the black box indicated with r represents the parallel channel due to the scattering of y around the regression line.

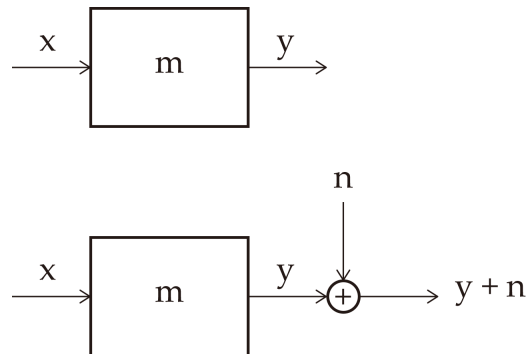


Figure 1. Flow chart of linear systems. **Upper panel:** Deterministic channel with multiplicative distortion (bias) m . **Lower panel:** Noisy channel with multiplicative bias and additive noise source.

Now, let us consider:

- The variance of the difference $mx - m_o x$, referred to as the “regression noise” power N_m .
- The variance of the difference between the values not lying on Equation (3) ($r \neq 1$), and those lying on it ($r = 1$), referred to as the “correlation noise” power N_r . This “noise” is due to the spread of y around the line given by Equation (3), due to the stochastic noise n .
- The variances s_x^2 and s_y^2 of x and y involved in Equation (3).

In case (a), we get the difference $(m - m_o)x$, therefore the variance (or power) of the values of y lying on the regression line is given by:

$$N_m = (m - m_o)^2 s_x^2. \quad (4)$$

Now, we define the “regression noise-to-signal power ratio” R_m (NSRm) between the output and the input:

$$R_m = \frac{N_m}{s_x^2} = (m - m_o)^2 \quad (5)$$

In case (b), the fraction of the variance s_y^2 due to the values of y not lying on the regression line is given by:

$$N_r = (1 - r^2) s_y^2. \quad (6)$$

The parameter r^2 is called the coefficient of determination and it gives the fraction of the variance of y explained by the regression line $r^2 s_y^2$ [19]. How-

ever, this variance is correlated with the slope m because $r^2 s_y^2$ is related to m according to:

$$r^2 s_y^2 = m^2 s_x^2. \quad (7)$$

Figure 2 shows the flow chart of variances.

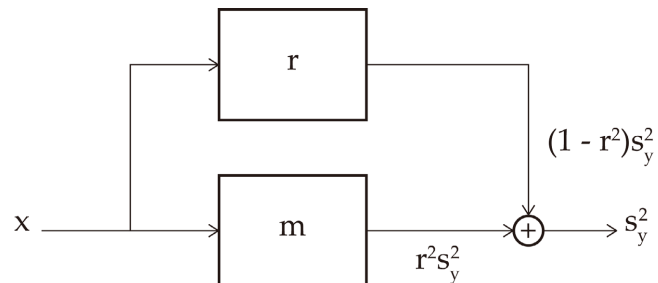


Figure 2. Flow chart of variances: $r^2 s_y^2$ is the output variance of the values lying on the regression line; $(1 - r^2) s_y^2$ is the output variance due to the values of y not lying on the regression line.

Therefore, inserting Equation (7) in Equation (6), we get the “correlation noise-to-signal power ratio”, R_r (NSRr):

$$R_r = \frac{N_r}{s_x^2} = \frac{1 - r^2}{r^2} m^2 \quad (8)$$

Since the two noise variances (powers) are additive, the total noise-to-signal (NSR) R of the channel shown in **Figure 2** is given by:

$$R = \frac{N_m}{s_x^2} + \frac{N_r}{s_x^2} = R_m + R_r \quad (9)$$

Therefore, R depends only on the two parameters m and r of the regression line, independently of their sign, explicitly given by:

$$R = (m - m_o)^2 + \frac{1 - r^2}{r^2} m^2. \quad (10)$$

Finally, the signal-to-noise ratio (SNR) γ , is given by:

$$\gamma = \frac{1}{R} = \frac{1}{(m - m_o)^2 + \frac{1 - r^2}{r^2} m^2}. \quad (11)$$

$$\Gamma(\text{dB}) = 10 \times \log_{10} \gamma \quad (12)$$

Of course, we expect that no real channel can yield $r = \pm 1$ and $m = m_o$, in which case $\gamma = \infty$, a channel purely deterministic. In experimental scatterplots, very likely, $|r| < 1$, $m \neq m_o$. It should be noted that the SNR can be considered an indicator of the degree of deterministic nature of a channel: the higher the SNR, the greater the “degree” of the channel’s deterministic nature.

In conclusion, the slope m measures the multiplicative “distortion” of the dependent variable y with respect to the independent variable x in the deter-

ministic channel; the correlation coefficient r measures how “precise” the best linear fit is.

Finally, notice the more direct and insightful analysis that can be achieved by using the NSR instead of the more common SNR because, in Equations (9) and (10), the single channel NSRs simply add together. This makes easier to study, for example, which addend determines R , and thus γ , while this is by far less easy with Equation (11).

Figure 3 shows the signal-to-noise ratio drawn with varying multiplicative bias m and $r = 0.99$, by assuming $m_o = 1$ (blue lines), $m_o = 1.5$ (red lines); $m_o = 2$ (green lines). Notice that at $m = m_o$, $R_m = 0$ and $\frac{1}{R_m} = \infty$. At these points, R is determined only by R_r , $R = R_r$, dashed lines. The lower continuous lines refer the full signal-to-noise ratio Γ .

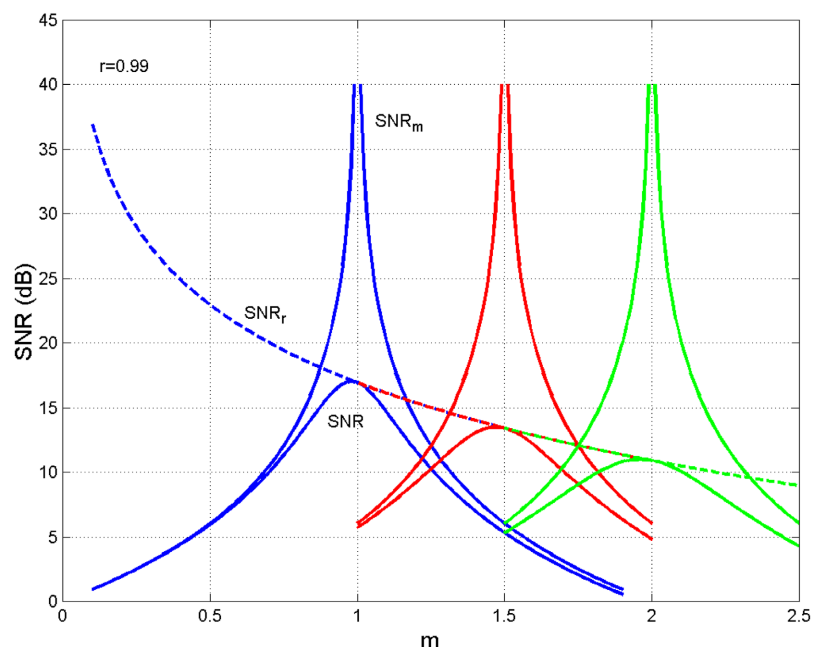


Figure 3. Signal-to-noise ratio drawn as a function of m , by assuming $r = 0.99$ for all, and $m_o = 1$ (blue lines), $m_o = 1.5$ (red lines); $m_o = 2$ (green lines). Notice that at $m = m_o$, $R_m = 0$ and $\frac{1}{R_m} = \infty$, see line labelled SNR_m . At these points, $R = R_r$, see line labelled SNR_r , dashed lines. The lower continuous lines refer to the full signal-to-noise ratio Γ .

Figure 4 shows the full signal-to-noise ratio Γ as a function of m and r . We can notice that as $|r| \rightarrow 1$, $\Gamma \rightarrow \infty$, at $m \cong m_o$. In this region, Γ changes very rapidly as small variations in r give very large variations in Γ , for any $m \cong m_o$.

The choice of using the NSR instead of the more common SNR also leads to a useful graphical representation of Equation (9) that can guide analysis and design, first used in my deep-space communication studies [20], as shown in the next section.

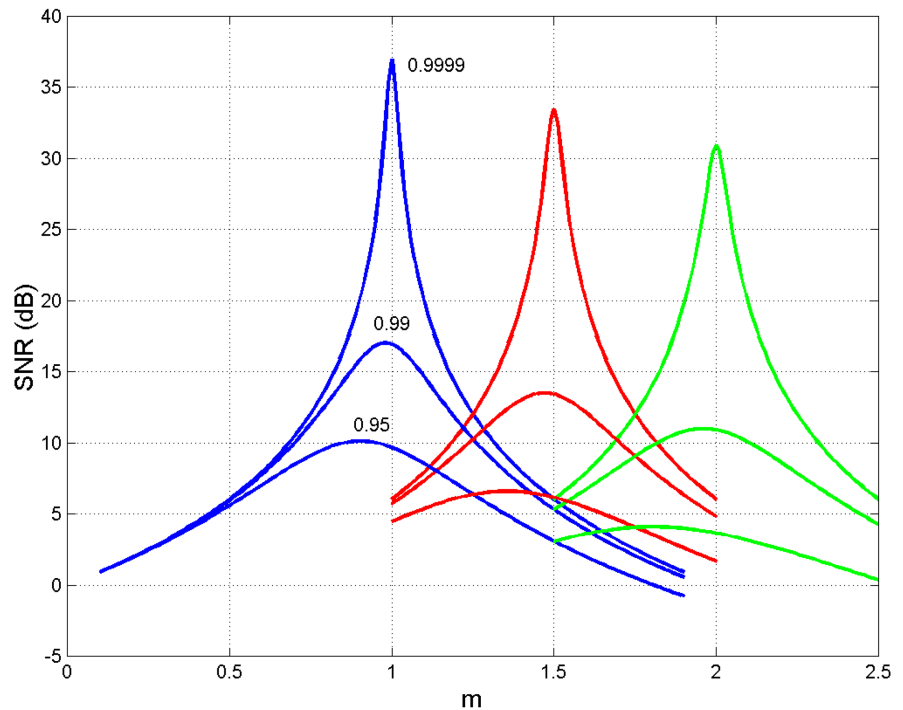


Figure 4. Γ versus m and r ; as $r \rightarrow 1$, $\Gamma \rightarrow \infty$, at $m \cong m_o = 1; 1.5; 2.0$. In this region, Γ changes vary rapidly as small variations in r give very large variations in Γ , for any m_o .

3. Geometrical Representation of the Noise-to-Signal Power Ratio

As discussed in [20] for deep-space radio link design, among other features not of interest here, adopting the noise-to-signal power ratio instead of the more common signal-to-noise ratio allows a geometrical representation, which immediately shows how R_r and R_m , through their square roots, contribute to the total R .

In fact, we can represent Equation (10) graphically by considering the two orthogonal variables:

$$X = \sqrt{R_m} \quad (13)$$

$$Y = \sqrt{R_r} \quad (14)$$

By setting the noise-to-signal ratio $R = R_o$, being $R_o = 1/\gamma_o$ a constant, X and Y trace an arch of the circle with radius $\sqrt{R_o} = \sqrt{R_m + R_r}$ in the first Cartesian quadrant. All points inside the circle correspond to $R < R_o$ ($\gamma > \gamma_o$); the origin of the axes corresponds to $R = 0$ of the purely deterministic channel, $m = m_o$ and $r = 1$. The signal-to-noise power ratio $\gamma = 1/R$ becomes infinite at the origin and decreases as the radius of the circle increases. **Figure 5** shows a universal chart with lines of equal signal-to-noise ratio Γ .

Channels can be combined and compared in various ways; therefore, in the following sections, I will examine the most common combinations and make some comparisons.

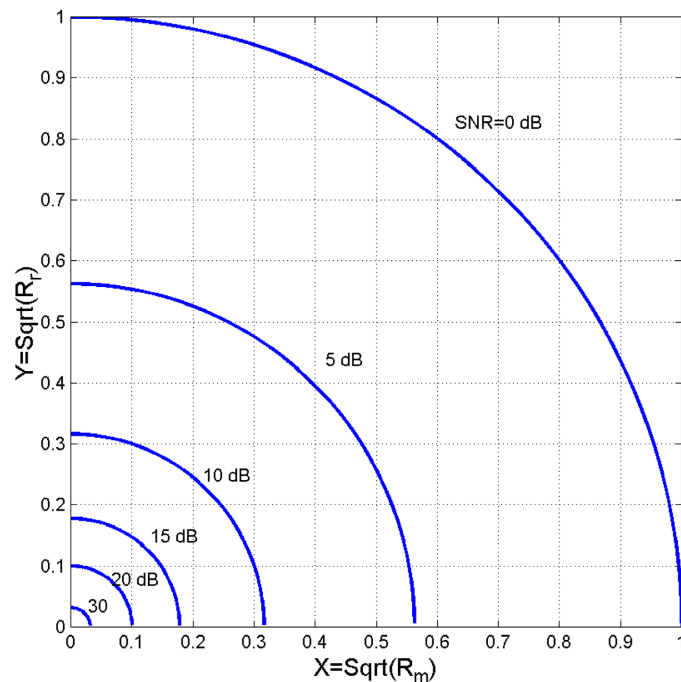


Figure 5. Lines of equal signal-to-noise ratio drawn in the universal chart of coordinates $X = \sqrt{R_m}$ and $Y = \sqrt{R_r}$.

4. Cross Channel

Let us study how the output variable y_k of channel k relates to the output variable y_j of another similar channel j for the same input $x_k = x_j = x$. This new channel is termed “cross channel”.

Let us consider the scatterplot k , referred to the couple x_k, y_k and the scatterplot j , referred to the couple x_j, y_j , in which the independent variable x and the dependent variable y are linked by the following regression lines in what we can define a “deterministic channel”:

$$y_k = m_k x_k \quad (15)$$

$$y_j = m_j x_j \quad (16)$$

As discussed in Section 2, Equations (15) and (16) do not give the full relationship between the two variables because they link only conditional average values, measured by the slopes m_k and m_j . According to Equation (3), we can write more general linear relationships by considering the scattering of the data, always present in experiments, modelled by zero-mean noise sources n_k and n_j .

$$y_k = m_k x_k + n_k \quad (17)$$

$$y_j = m_j x_j + n_j \quad (18)$$

Now, we can develop a series of interesting investigations on these equations. By eliminating x , we can compare the dependent variable y_j of Equation (18) to the dependent variable y_k of Equation (17) for $x_k = x_j = x$. In doing so, we can find the regression line and the correlation coefficient of the new scatterplot

linking y_j to y_k without the availability of the scatterplot itself, assuming the same value of the independent variable x .

By eliminating x between Equation (17) and Equation (18), we get:

$$y_j = \frac{m_j}{m_k} y_k - \frac{m_j}{m_k} n_k + n_j. \quad (19)$$

Compared to the new independent variable y_k , the slope m_{kj} of the regression line is given by:

$$m_{kj} = m_j / m_k. \quad (20)$$

Because the two noise sources are additive, the total noise is given by:

$$n_{kj} = -\frac{m_j}{m_k} n_k + n_j = -m_{kj} n_k + n_j. \quad (21)$$

Figure 6 shows the flow chart describing the cross-channel.

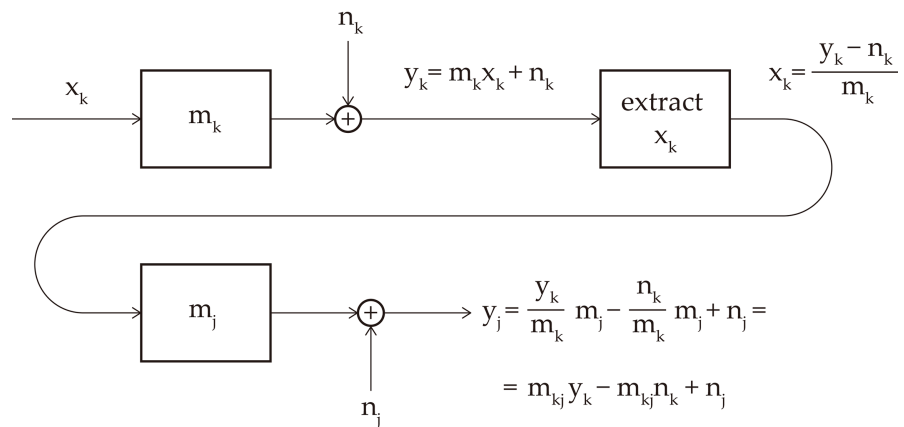


Figure 6. Chart describing the cross-channel signal flow.

Now, defining the expected (deterministic) ratio between the two slopes $m_{o,kj} = m_{o,j} / m_{o,k}$, we get R_m of the cross channel:

$$R_m = (m_{kj} - m_{o,kj})^2. \quad (22)$$

To proceed to find R_r , it is necessary to determine the correlation coefficient r_{kj} between $-m_{kj} n_k$ and n_j because in Equation (21), the two noise sources might be correlated, with unknown correlation coefficient r_{kj} . **Appendix B** shows that:

$$r_{kj} = \cos \left| \arccos(r_j) - \arccos(r_k) \right| \quad (23)$$

Therefore, R_r of the cross channel is:

$$R_r = \frac{1 - r_{kj}^2}{r_{kj}^2} m_{kj}^2. \quad (24)$$

In conclusion, we can determine the slope and the correlation coefficient of the scatterplot between y_j and y_k for the same value of the independent variable x . Now, the availability of this scatterplot is experimentally very rare because it is

unlikely to find values of y_k and y_j for exactly the same value of x , therefore cross channels can reveal relationships very difficult to discover experimentally.

In the next section, we study how the output of a deterministic channel relates to the output of its stochastic version.

5. Stochastic Channel versus Deterministic Channel

We compare a deterministic channel k (slope m) with a stochastic channel j derived from channel k by adding noise (slope m , correlation coefficient r). This is a particular cross channel, therefore, from the theory of Section 4, we get:

$$y_k = m_k x_k, \quad (25)$$

$$y_j = m_k x_k + n_k, \quad (26)$$

$$m_{kj} = m_k / m_k = 1, \quad (27)$$

$$R_m = (m_{kj} - 1)^2 = 0, \quad (28)$$

$$r_{kj} = \cos \left| \arccos(r_j) - \arccos(1) \right| = \cos \left| \arccos(r_j) \right| = r_j = r \quad (29)$$

$$R = R_r = \frac{1 - r^2}{r^2} \quad (30)$$

In conclusion, in transforming the deterministic channel into its stochastic version, only the correlation noise is present, therefore the SNR is given by:

$$\gamma = \frac{r^2}{1 - r^2}. \quad (31)$$

Equation (31) coincides with the ratio between the variance explained by the regression line (proportional to r^2) and the variance due to the scattering (correlation noise), proportional to $1 - r^2$.

The development of this section confirms that the SNR can be considered—in any case—an indicator of the degree of deterministic nature of a channel: the higher the SNR, the greater the “degree” of the channel’s deterministic nature.

So far, we have considered single channels. In the following sections, we examine the series and parallel connection of channels and determine the signal-to-noise ratio of the overall channel.

6. Parallel Connection of Channels

Let us consider the parallel connection of the channels, **Figure 7**. The parallel connection gives a single channel with the following input-output relationship:

$$y = y_k + y_j = m_k x_k + n_k + m_j x_j + n_j, \quad (32)$$

Now, since $x_k = x_j = x$, Equation (32) is equivalent to the single channel

$$y = (m_k + m_j) x_k + n_k + n_j \quad (33)$$

Therefore, Equation (33) coincides with Equation (3) in which:

$$m = m_k + m_j \quad (34)$$

$$n = n_k + n_j \quad (35)$$

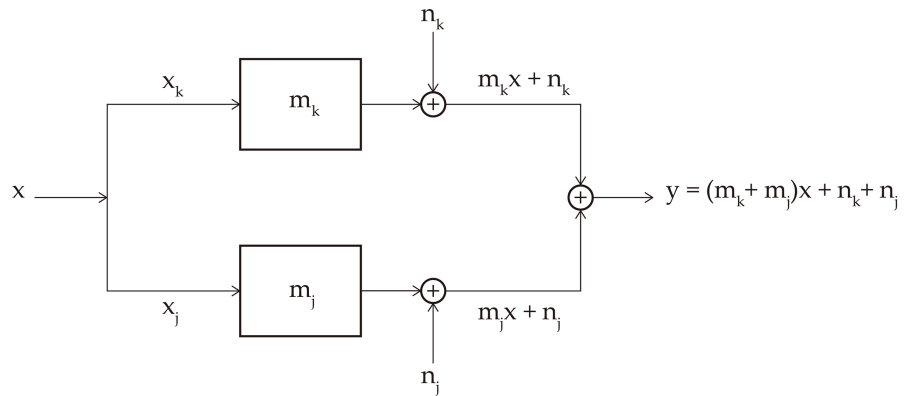


Figure 7. Parallel connection of channels.

Let us calculate the correlation coefficient of this channel. From Equation (7)

$$r^2 = m^2 \frac{s_x^2}{s_y^2}. \quad (36)$$

Therefore, after a few mathematical steps (**Appendix C**), we get:

$$r = \sqrt{\frac{(m_k + m_j)^2}{\frac{m_k^2}{r_k^2} + \frac{m_j^2}{r_j^2} + 2r_{kj} \frac{m_k}{r_k} \frac{m_j}{r_j}}} \quad (37)$$

Notice that if $m_k = m_j$ then

$$r = \sqrt{\frac{4}{\frac{1}{r_k^2} + \frac{1}{r_j^2} + \frac{2r_{kj}}{r_k r_j}}} \quad (38)$$

Figure 8 shows how r varies as a function of r_k and r_j according to Equation (38). Notice that if $r_k = r_j = 1$, then $r = 1$.

In conclusion, the noise-to-signal ratios R_m and R_r are given, respectively, by Equation (5) and Equation (8), with m given by Equation (34) and r given by Equation (37). Of course, the theory of parallel channels can be extended to the parallel of $n > 2$ single channels by applying Equations (34) and (37) recursively.

7. Series Connection of Channels

In this section, we consider a channel made by a series of single channels, **Figure 9**. The “experimental” slope of the overall channel is given by:

$$m = m_1 m_2 m_3. \quad (39)$$

Therefore, in a series of n channels, the expected overall slope is given by;

$$m_o = \prod_{i=1}^n m_{o,i} \quad (40)$$

Therefore, the noise-to-signal regression noise R_m is given by:

$$R_m = (m - m_o)^2 = \left(\prod_{i=1}^n m_i - \prod_{i=1}^n m_{o,i} \right)^2 \quad (41)$$

From **Figure 9**, it is evident that the output noise of a preceding channel produces additive noise at the output of the next channels. Let us calculate the correlation noise-to-signal ratio R_r at the output of the overall channel.

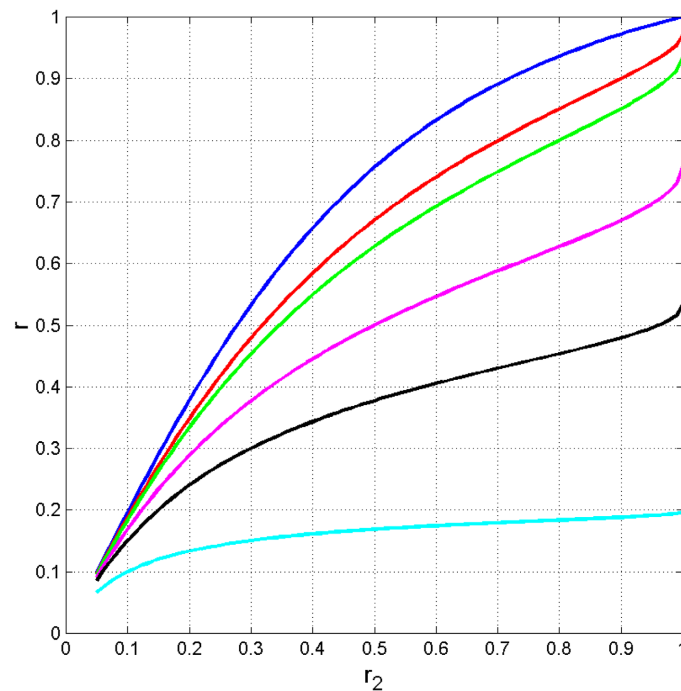


Figure 8. Curves showing how r varies as a function of $r_k = r_2$ versus $r_j = r_1$, by assuming $m_k = m_j$, according to Equation (38). Cyan line $r_1 = 0.1$; black line $r_1 = 0.3$; magenta line $r_1 = 0.5$; green $r_1 = 0.8$; red line $r_1 = 0.9$; blue line $r_1 = 1$.

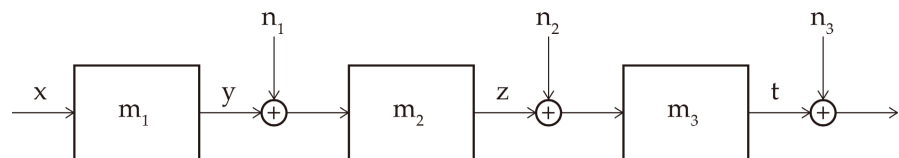


Figure 9. Series connection of channels.

Theorem. The correlation noise-to-signal ratio $R = R_r$ of n linear channels in series, each characterized by the correlation noise-to-signal ratio R_i , is given by:

$$R = \sum_{i=1}^n R_i. \quad (42)$$

Proof. Let the three linear input-output relationships of the isolated channels of **Figure 9** (i.e., before connecting them in series) given by:

$$y = m_1 x + n_1, \quad (43)$$

$$z = m_2 y + n_2, \quad (44)$$

$$t = m_3 z + n_3. \quad (45)$$

Let s_x^2 , s_y^2 , s_z^2 the variances (power) of the input variables to the single

channels, and let $N_1 = N_{1r}$, $N_2 = N_{2r}$, $N_3 = N_{3r}$ the variances (power) of the zero-mean noise n_1 , n_2 , n_3 , then the NSRs of the isolated channels are given by:

$$R_1 = \frac{N_1}{s_y^2} = \frac{N_1}{m_1^2 s_x^2} \quad (46)$$

$$R_2 = \frac{N_2}{s_z^2} = \frac{N_2}{m_2^2 s_y^2} \quad (47)$$

$$R_3 = \frac{N_3}{s_t^2} = \frac{N_3}{m_3^2 s_z^2} \quad (48)$$

When the first two channels are connected in series, the input to the second channel must also include the output noise of the first channel, therefore we get the modified output variable z :

$$z = m_2 y = m_2 (m_1 x + n_1) + n_2 = m_2 m_1 x + m_2 n_1 + n_2. \quad (49)$$

In Equation (49), $m_2 m_1 x$ is the output “signal” and $m_2 n_1 + n_2$ is the output noise, therefore, the NSR at the output of the second channel is:

$$R = \frac{m_2^2 N_1 + N_2}{m_2^2 m_1^2 s_x^2} = \frac{m_2^2 N_1}{m_2^2 m_1^2 s_x^2} + \frac{N_2}{m_2^2 m_1^2 s_x^2} = \frac{N_1}{m_1^2 s_x^2} + \frac{N_2}{m_2^2 s_y^2} = R_1 + R_2 \quad (50)$$

Now, for three channels in series, it is sufficient to consider R given by Equation (50) as the input NSR to the third single channel to obtain the final NSR and prove Equation (42):

$$R = R_1 + R_2 + R_3. \quad (51)$$

□

Finally, notice that R of Equation (51) is proportional to the mean $\langle R_i \rangle$:

$$R = n \sum_{i=1}^n R_i / n = n \langle R_i \rangle. \quad (52)$$

In other words, the series connection of channels averages the single correlation noise-to-signal ratios R_i .

8. Summary and Conclusions

I have summarized the results of a theory regarding linear regression developed in series of papers [1]-[17]. The linear relationship between the dependent variable y and the independent variable x is interpretable as describing the input-output characteristic of a linear system, or a linear channel.

The conviction that the theory is applicable to any scientific discipline—provided a linear input-output relationship exists—has led me to summarize the main results in a general form that the readers can easily apply to their specific problem.

Defining the regression noise-to-signal power ratio (NSR) R_m and the correlation noise-to-signal power ratio, R_r , I have shown that the overall noise-to-signal ratio R is the sum of the two. I have highlighted the more direct and insightful analysis achievable by using the NSR instead of the more common signal-to-noise ratio (SNR) because it makes easier to study which addend determines

R and allows a useful graphical representation.

I have underlined that the SNR can be considered an indicator of the degree of deterministic nature of a channel, because the higher the SNR, the greater the “degree” of the channel’s deterministic nature.

In the paper, I have considered various connections of single channels (single, cross, deterministic, parallel, series), whose main results are summarized in **Table 1**.

Table 1. Channels configuration parameters. For the meaning of the mathematical symbols, see the main text and **Appendix A**.

Channels configuration	R_m	R_r	Correlation coefficient	Slope
Single	$(m - m_o)^2$	$\frac{1 - r^2}{r^2} m^2$	-	-
Cross	$(m_{kj} - m_{o,kj})^2$	$\frac{1 - r_{kj}^2}{r_{kj}^2} m_{kj}^2$	$\cos \arcsin(r_j) - \arcsin(r_k) $	m_j / m_k
Parallel	$(m_{kj} - m_{o,kj})^2$	$\frac{1 - r_{kj}^2}{r_{kj}^2} m_{kj}^2$	$\sqrt{\frac{(m_k + m_j)^2}{\frac{m_k^2}{r_k^2} + \frac{m_j^2}{r_j^2} + 2r_{kj} \frac{m_k}{r_k} \frac{m_j}{r_j}}}$	$m_k + m_j$
Series	$\left(\prod_{i=1}^n m_i - \prod_{i=1}^n m_{o,i} \right)^2$	$\sum_{i=1}^n R_i$	-	$\prod_{i=1}^n m_i$

In conclusion, the theory and its mathematical results concerning the input-output characteristics of linear channels, described by experimental regression lines, should be useful in any scientific discipline in which the experimental problem is modelled linearly by regression lines.

Acknowledgements

I wish to thank Lucia Matricciani for drawing **Figure 1**, **Figure 2**, **Figure 6**, **Figure 7**, **Figure 9**, **Figure A1**, **Figure A2**.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Matricciani, E. (2019) Deep Language Statistics of Italian Throughout Seven Centuries of Literature and Empirical Connections with Miller’s 7 ∓ 2 Law and Short-Term Memory. *Open Journal of Statistics*, **9**, 373-406.
<https://doi.org/10.4236/ojs.2019.93026>
- [2] Matricciani, E. (2020) A Statistical Theory of Language Translation Based on Communication Theory. *Open Journal of Statistics*, **10**, 936-997.
<https://doi.org/10.4236/ojs.2020.106055>

- [3] Matricciani, E. (2022) Linguistic Mathematical Relationships Saved or Lost in Translating Texts: Extension of the Statistical Theory of Translation and Its Application to the New Testament. *Information*, **13**, Article No. 20. <https://doi.org/10.3390/info13010020>
- [4] Matricciani, E. (2022) Multiple Communication Channels in Literary Texts. *Open Journal of Statistics*, **12**, 486-520. <https://doi.org/10.4236/ojs.2022.124030>
- [5] Matricciani, E. (2023) Capacity of Linguistic Communication Channels in Literary Texts: Application to Charles Dickens' Novels. *Information*, **14**, Article No. 68. <https://doi.org/10.3390/info14020068>
- [6] Matricciani, E. (2023) Readability Indices Do Not Say It All on a Text Readability. *Analytics*, **2**, 296-314. <https://doi.org/10.3390/analytics2020016>
- [7] Matricciani, E. (2023) Readability across Time and Languages: The Case of Matthew's Gospel Translations. *AppliedMath*, **3**, 497-509. <https://doi.org/10.3390/appliedmath3020026>
- [8] Matricciani, E. (2023) Linguistic Communication Channels Reveal Connections between Texts: The New Testament and Greek Literature. *Information*, **14**, Article No. 405. <https://doi.org/10.3390/info14070405>
- [9] Matricciani, E. (2023) Short-Term Memory Capacity across Time and Language Estimated from Ancient and Modern Literary Texts. Study-Case: New Testament Translations. *Open Journal of Statistics*, **13**, 379-403. <https://doi.org/10.4236/ojs.2023.133019>
- [10] Matricciani, E. (2024) Is Short-Term Memory Made of Two Processing Units? Clues from Italian and English Literatures down Several Centuries. *Information*, **15**, Article No. 6. <https://doi.org/10.3390/info15010006>
- [11] Matricciani, E. (2024) A Mathematical Structure Underlying Sentences and Its Connection with Short-Term Memory. *AppliedMath*, **4**, 120-142. <https://doi.org/10.3390/appliedmath4010007>
- [12] Matricciani, E. (2024) Multi-Dimensional Data Analysis of Deep Language in J.R.R. Tolkien and C.S. Lewis Reveals Tight Mathematical Connections. *AppliedMath*, **4**, 927-949. <https://doi.org/10.3390/appliedmath4030050>
- [13] Matricciani, E. (2025) A Mathematical Analysis of Texts of Greek Classical Literature and Their Connections. *Open Journal of Statistics*, **15**, 1-34. <https://doi.org/10.4236/ojs.2025.151001>
- [14] Matricciani, E. (2025) Translation Can Distort the Linguistic Parameters of Source Texts Written in Inflected Language: Multidimensional Mathematical Analysis of "the Betrothed", a Translation in English of "I Promessi Sposi" by A. Manzoni. *AppliedMath*, **5**, Article No. 24. <https://doi.org/10.3390/appliedmath5010024>
- [15] Matricciani, E. (2025) Domestication of Source Text in Literary Translation Prevails over Foreignization. *Analytics*, **4**, Article No. 17. <https://doi.org/10.3390/analytics4030017>
- [16] Matricciani, E. (2025) Revealing Short-Term Memory Communication Channels Embedded in Alphabetical Texts: Theory and Experiments. *Information*, **16**, Article No. 847. <https://doi.org/10.3390/info16100847>
- [17] Matricciani, E. (2025) Equivalent Processors Modelling the Short-Term Memory. *International Journal of Pure and Applied Mathematics Research*, **5**, 1-28. <https://doi.org/10.51483/ijpamr.5.2.2025.1-28>
- [18] De Caro, L., Giannini, C., Lassandro, R., Scattarella, F., Sibillano, T., Matricciani, E., et al. (2019) X-Ray Dating of Ancient Linen Fabrics. *Heritage*, **2**, 2763-2783.

<https://doi.org/10.3390/heritage2040171>

- [19] Papoulis, A. (1990) Probability & Statistics. Prentice Hall.
- [20] Matricciani, E. (2009) An Optimum Design of Deep-Space Downlinks Affected by Tropospheric Attenuation. *International Journal of Satellite Communications and Networking*, **27**, 312-329. <https://doi.org/10.1002/sat.942>
- [21] Lindgren, B.W. (1968) Statistical Theory. 2nd Edition, MacMillan Company.

Appendix A

Table A1. lists of the mathematical symbols used in the article with their meaning.

Symbol	Definition
m	Slope of regression line
m_o	Expected slope
m_{kj}	Slope in combination of single channels
m_{kj}	Correlation coefficient in combination of single channels
r	Correlation coefficient of linear variables
r^2	Coefficient of determination
s	Standard deviation
s^2	Variance
N_m	Regression noise power
N_r	Correlation noise power
R	Noise-to-signal overall power ratio
R_m	Regression noise-to-signal power ratio
R_r	Correlation noise-to-signal power ratio
$\check{y}$	Dependent variable with intercept zero
γ	Signal-to-noise ratio (linear)
Γ	Signal-to-noise ratio (dB)
$\langle \rangle$	Mean value

Appendix B

Consider the stochastic variables x and y with variances s_x^2 and s_y^2 respectively. Consider two vectors in the plane whose lengths are proportional to s_x and s_y and separated by an angle θ whose cosine is given by the correlation coefficient r (**Figure A1**). The length of the diagonal of the parallelogram determined by them can be computed using the law of cosines ([21], p. 127):

$$d^2 = s_x^2 + s_y^2 - 2s_x s_y \cos(\pi - \theta) = s_x^2 + s_y^2 + 2r s_x s_y \quad (\text{A1})$$

Now Equation (A1) is just the variance of the stochastic variable $z = x + y$, therefore the length of the diagonal d is the standard deviation of z . Thus, the addition of the stochastic variables x and y corresponds to vector addition of the vectors representing their standard deviations.

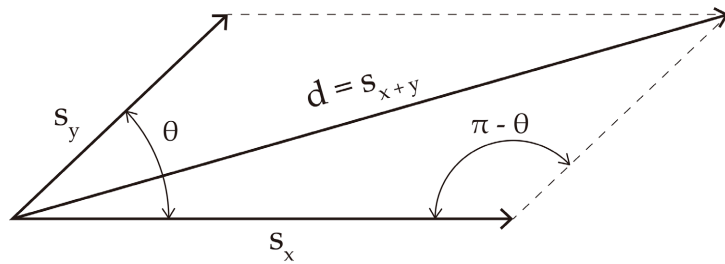


Figure A1. Geometrical representation of standard deviations ([21], p. 127). The standard deviations can be represented by two vectors whose lengths are proportional to s_x and s_y respectively and separated by an angle θ whose cosine is given by the correlation coefficient r , according to the law of cosines.

Let us now apply this representation to the relationships of the problem in this article:

$$y_k = m_k x + n_k \quad (\text{A2})$$

$$y_j = m_j x + n_j \quad (\text{A3})$$

We get

$$s_{yk}^2 = m_k^2 s_x^2 + s_{nk}^2 + 2r_k m_k s_x s_{nk} \quad (\text{A4})$$

$$s_{yj}^2 = m_j^2 s_x^2 + s_{nj}^2 + 2r_j m_j s_x s_{nj} \quad (\text{A5})$$

We wish to calculate the correlation coefficient r_{kj} between the two noise sources, here rewritten by changing the negative sign into positive without any impact on the calculations when dealing with noise:

$$n_{kj} = m_{kj} n_k + n_j \quad (\text{A6})$$

Figure A2 shows the vector representations of the standard deviations of Equations (A4) and (A5).

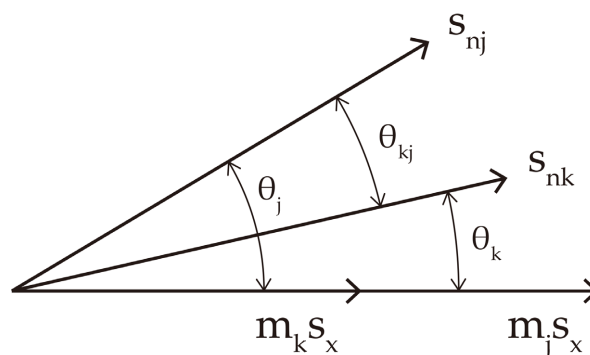


Figure A2. Vector representations of the standard deviations of Equations (A4) and (A5) to find the correlation coefficient r_{kj} between $m_{kj} n_k$ and n_j .

Now r_{kj} is the cosine of the angle θ_{kj} between the vectors s_{nj} and s_{nk} , therefore:

$$r_{kj} = \cos(\theta_{kj}) = \cos|\theta_j - \theta_k| \quad (\text{A7})$$

Now, since

$$\theta_k = \arccos(r_k) \quad (\text{A8})$$

$$\theta_j = \arccos(r_j) \quad (\text{A9})$$

We get the final relationship, Equation (23):

$$r_{kj} = \cos \left| \arccos(r_j) - \arccos(r_k) \right| \quad (\text{A10})$$

Appendix C

In Equation (36), the variance, s_y^2 , of the two correlated stochastic variables in Equation (32), is given by Equation (A1), here written for our problem:

$$s_y^2 = s_{y,k}^2 + s_{y,j}^2 + 2r_{kj}s_{y,k}s_{y,j} \quad (\text{A11})$$

In Equation (A11), r_{kj} is given by Equation (23). Therefore:

$$r^2 = m^2 \frac{s_x^2}{s_y^2} = m^2 \frac{s_x^2}{s_{y,k}^2 + s_{y,j}^2 + 2r_{kj}s_{y,k}s_{y,j}} \quad (\text{A12})$$

$$r^2 = m^2 \frac{s_x^2}{\frac{m_k^2}{r_k^2} s_x^2 + \frac{m_j^2}{r_j^2} s_x^2 + 2r_{kj}s_x^2 \sqrt{\frac{m_k^2}{r_k^2} \frac{m_j^2}{r_j^2}}} \quad (\text{A13})$$

$$r^2 = m^2 \frac{1}{\frac{m_k^2}{r_k^2} + \frac{m_j^2}{r_j^2} + 2r_{kj} \sqrt{\frac{m_k^2}{r_k^2} \frac{m_j^2}{r_j^2}}} \quad (\text{A14})$$

From Equation (A14), we get Equation (37).