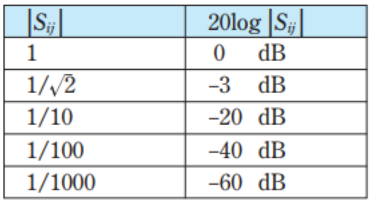
Asu Naz Yıldırım

Homework 1

In this homework, the aim is to understand how 11 antenna design parameters affect S_{11} parameters. We try to find an answer to whether Principal Component Analysis (PCA) and/or linear regression can be used in designing antennas.

S_{11}, also referred to as reflection coefficient or return loss, shows the amount of power reflected from the antenna [1]. This parameter is a complex number [2]. Although it has no dimension, "dB" is usually used in conjunction with logarithm with base 10 to describe its magnitude [3].

Table 1. Typical S-Parameter Values [3].



$\lvert S_{ij} \rvert$	$20\log \lvert S_{ij} \rvert$
1	0 dB
$1/\sqrt{2}$	–3 dB
1/10	–20 dB
1/100	–40 dB
1/1000	–60 dB

Lower amounts of reflected power are indicated by greater magnitudes of return loss in absolute sense [4]. Ideally, the S_{11} value should be as close to -∞ as possible [5]. In practice, -20 dB is seen as a very good S_{11} value [5]. Many radio frequency engineers consider -10 dB as the acceptable limit for good matching [5]. Values above -10 dB are generally viewed as mismatched [5].

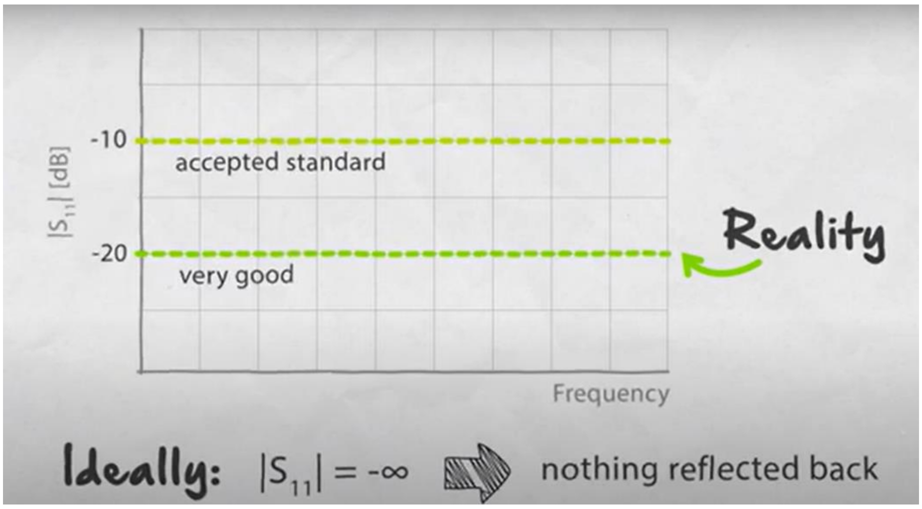


Figure 1. Standard S_{11} limits [5].

Dimensionality Reduction with PCA:

- Can we reduce the complexity of the design space by using Principal Component Analysis (PCA) to identify key parameters that most influence the S_{11} response of the antenna?

Principal Component Analysis (PCA) is an unsupervised machine learning technique which aims to reduce the dimensionality of the data while maintaining the most important relationship between the variables [6]. Therefore, PCA does not require any prior knowledge of the target variable [6]. It is used exploratory data analysis [7]. Additionally, one should keep in mind that it assumes linear relationship between variables [8].

In our problem, if PCA is applied to design space, we obtain information about the relationship between the design parameters rather than how they influence S_{11} response. However, thanks to PCA, after we detect the directions (i.e. principal components) that contains the maximum variance in the data set, we now have new representations (i.e. latent variables) in the transformed space which needs to be interpreted as feature index. If we can obtain an interpretable result in latent variables with the highest variance, we can reduce the complexity of the design space to detect the most correlated design parameters. Unfortunately, it does not guarantee how much these latent variables influence S_{11} response of the antenna.

To apply PCA, first I standardized the data with zscore function to avoid any scaling issues. After pca function in matlab is applied on the scaled data:

Each column of coeff matrix corresponds to principal components (i.e. eigenvectors) starting with one with the highest eigenvalue to the lowest one.

Each row of coeff matrix corresponds to the coefficients for the design variables.

Each row of explained matrix is the percentage of the total variance for each principal component.

To illustrate, we consider the first principal component, it has the highest variance of 20% of the total variance. Its coefficients are given in Table 1.

Table 1. PC 1 coefficients.

	PC 1
length of patch	-0.1013
width of patch	0.6242
height of patch	0.0711
height of substrate	0.6241
height of solar resist layer	-0.0210
radius of the probe	0.0260
c_pas	-0.0542
c_antipad	-0.0187
c_probe	0.0392
dielectric constant of substrate	0.4451
dielectric constant of solder resist layer	0.0384

Since the norm of this vector is unity, the coefficients are between 1 and -1. A positive coefficient indicates that the corresponding design parameter positively contributes to the loading. On the other hand, a negative coefficient shows that the corresponding design parameter negatively contributes to the loading. To illustrate, if we can interpret PC 1 as a possible feature index, a higher score in this principal component direction shows an antenna has high patch width and high substrate height and low patch length. The magnitude of the coefficient shows how strongly the contribution exists. However, to set a standard to describe which design parameter contributes meaningfully, one can say that if the coefficient is higher than 0.5 or lower than -0.5, the corresponding design parameter is taken into account. As an example, For PC 1, the width of patch and the height of substrate are the main design parameters. Similar reasoning can be applied to all principal components.

- How much of the total variance in the design parameter space can be explained by the principal components?

Each row of explained matrix is the percentage of the total variance for each principal component.

PC 1: 20.7153%, PC 2: 11.0708%, PC 3: 10.0034%, PC 4: 9.6072%, PC 5: 9.1871%, PC 6: 8.8458%, PC 7: 8.5659%, PC 8: 8.1651%, PC 9: 7.3399%, PC 10: 5.8303%,

PC 11: 0.6692%

Although PC 1 has the largest variance of 20%, from PC 2 to PC 9 each principal component has approximately the same contribution of 10% to the total variance. It seems as if the data is randomly generated. To check, let's generate data (385x11) with randomly normal distributed.

- What insights can we draw from the PCA regarding the relationship between geometry and electromagnetic behavior?

Except the first and last principal components, two graphs are very similar. We may say that the design parameters are not that correlated. Therefore, it is very hard to conclude how these design parameters influence S_{11} by performing PCA. Indeed, since we do not consider the output variable (S_{11} in this case), one cannot capture the direct relationship between geometry and S_{11}.

However, if one still wants to reduce dimension, the code above can be used to determine the minimum number of principal components needed to obtain a specific cumulative explained variance threshold. This threshold is set at 80% but it depends on the application and may vary. Then, new representations of the design parameters can be obtained up to the selected principal component according to threshold.

Regression Modeling for S_{11}:

- Given that S_{11} parameters are evaluated at 201 frequency points, predicting them simultaneously through multitarget regression may be computationally intensive and beyond the scope of the content covered so far. Instead, can we simplify the regression task by focusing on predicting S_{11} at a few key frequency points? For instance, selecting frequencies of interest where the behavior is most critical (e.g., resonance frequencies) can reduce the complexity of the task.

Since S_{11} parameters depend on the frequency, output data composes of multiple columns. To simplify the task, instead of multitarget regression, if we focus on S_{11} at a few key frequency points, we will ignore how frequency and S_{11} parameters are related. However, we may think that if we can detect the frequency points beforehand, these points would be where we want to design antennas. Therefore, the task can be reduced to design antennas at the selected frequencies where antenna behavior is the most important. In this case, another criterion appears: what are these key frequency points?

This can be done in several ways as we will describe below. After detecting these frequencies, we will perform linear regression to understand the relation between the design and S_{11} parameter at the selected frequencies.

To visualize the data, I plotted S_{11} vs frequency both in terms of norm and decibel. However, selecting frequencies in terms of decibel is more meaningful because antenna design standard is reported based on dB scale [3] [5].

One method to identify key frequency points is to check the variance of S_{11} at each frequency point. The frequency where S_{11} has the maximum variance may be considered as the critical point and can be investigated.

Another method to identify key frequency points might be to detect the frequencies where local minima are observed. These frequencies can be interpreted as critical. I observed that including first and last frequencies mislead the detection of critical behavior if we look at the plot S_{11} in decibel vs. frequency. Therefore, I excluded these points. By using the bar plot, I visualized the "possible distribution" of the local minima. I found the most frequent frequency from the filtered data.

However, it was previously stated that -10 dB is considered as the acceptable limit. Therefore, we may eliminate antennas with inappropriate design and investigate the effect of design parameters on S_{11} in suitable antennas. If S_{11} of an antenna is -10 dB or lower for at least one frequency, we keep this antenna and the others are eliminated. Now, we can apply the same procedure as described above to identify critical points.

- How effective are linear regression models in predicting the real and imaginary components of S_{11} at these selected frequency points, based on the geometric parameters of the antenna design?
- What patterns emerge when linear regression is applied to individual frequency points, and do these patterns suggest any broader trends in the design space?

After performing the multiple linear regression between design parameters and S_{11} at the selected frequency points, I tried to interpret the models.

First, I checked the F-statistics. F-statistics is a significance test for the entire regression [9]. It shows if your linear regression model fits the data better than a model with no independent variables [10]. If we set the significance level (α) to 0.05 and p-value < 0.05, the regression becomes statistically significant [9]. Therefore, for all the selected frequency points, linear regression models are statistically significant.

To understand the nature of relationship and its statistical significance, coefficients along with its p-values should be interpreted [11]. While the coefficients show the relationship between the predictors and the response, their p-values indicate whether these relationships are statistically significant [11]. The intercept is the mean of the response when every predictor is equal to 0 [12]. The coefficient of a predictor shows how much the response changes when that predictor increases by 1 unit while keeping all other predictors constant [12].

Let us consider the model to predict the imaginary part of S_{11} at frequency 146. The intercept, the width of patch, the height of substrate, the radius of the prove and c_probe yield statistically significant coefficients among the estimated coefficients of the linear regression model to predict the imaginary part of S_{11} at frequency 146 if we set α = 0.05. According to this model, an increase in the radius of the probe results in substantial decrease in the imaginary part of S_{11}. On the other hand, an increase in the width of the patch results in a relatively small increase in the imaginary part of S_{11}.

[13]: To evaluate the model further, R^2 and adjusted R^2 are two key metrices. R^2 is the proportion of the variance in the response which is explained by the predictors. How well a model fits the data is measured by R^2, ranging from 0 (no predictive power) to 1 (perfect fit). However, R^2 always increases as more predictors are added to the model even if they don't actually improve the model. Therefore, adjusted R^2 addreses this problem in multiple linear regression. Adjusted R^2 takes the number of predictors into account and penalizes for adding irrelevant ones. Adjusted R^2 may go down if a new predictor is irrelavant to model. A higher adjusted R^2 means the model fits well with relevant predictors. A lower or negative value indicates adding more predictors may make the model worse by adding noise.

If we focus on the model to predict the imaginary part of S_{11} at frequency 146, R^2 is 0.308. This value shows that the model does not explain a large portion of the variance. There may be non-linear realtionship which a linear modeal cannot naturally capture. Since adjusted R^2 is slightly lower than R^2, some predictors may not make a significant contribution.

To investigate the effectivness of the model further, predicted values vs. observed values is plotted. If our model was perfect to predict the response, we would expect the points to align along a 45° line. However, it shows a substantial spread around that line. The predictions are not always close to the actual values. Therefore, this model does not seem suitable to the imaginary part of S_{11} at frequency 146.

Similar results (F-statistics, coefficients with p-values, R^2 and adjusted R^2, predicted values vs. observed values) are observed for the prediction of the imaginary part of S_{11} at frequency 145 as with frequency of 146. This is probably due to the fact that the frequencies are consecutive numbers. It is interesting to observe that the predicted values for the imaginary part are approximately bounded between -0.2 and 0.8. When the observed values are close to -1, the model sometimes predicts a positive value. Rather than aligning along the 45° line, the points tend to follow a flatter distribution. I checked if this trend is seen for other frequencies with high variance but far from frequency 146. I selected frequencies 155 and 71. They show a similar result. I also checked the model at frequency 114, which has the highest number of local minima. Although we obtain different coefficients with corresponding p-values if we set a linear model to predict the imaginary part of S_{11} at frequency 114, which is slightly far from 146, the plot of predicted values vs. observed values show the similar trend. To conclude, the multiple linear regression is not effective for this prediction of the imaginary part of S_{11} at these selected frequencies. Additionally, filtering the data based on the criterion of $S_{11} \leq -10\,\mathrm{dB}$ does not improve the model's performance either.

If the model to predict the real part of S_{11} is considered, a different result is obtained.

Let us focus on the model to predict the real part of S_{11} at frequency 146. The result of F-statistics shows that the model is again statistically significant. This time the significant predictors are the length of patch, the width of patch, the height of substrate, the radius of the probe, and the dielectric constant of substrate. R^2 and adjusted R^2 values are suprisingly high. R^2 is 0.8, meaning that the model explains 80% of the variance in the real part of S_{11}. This can be viewed as a good fit. Adjusted R^2 is 0.794, which is slightly lower than 0.8, shows the most of the predictors contribute meaningfully to the model. In the plot of predicted vs. observed values, an interesting pattern is observed. There seem to be a cluster around observed and predicted values close to -1. In this region, values appear to be more in line. However, as the observed value increases, we observe a spread in the predictions. We may say that the model's accuracy changes according to the ranges of the observed values. Additionally, there exist a gap. There are no predicted values between 0.2 and -0.2. It may be possible that there are two different distributions according to the ranges that we try to predict with a single model. Another possibility is that despite a high R^2 and adjusted R^2, there are non-linear relationships that should be considered to build the model. Similar results are observed for the prediction of the imaginary part of S_{11} at different frequencies. Again, the model's performance is not enhanced by filtering the data using the $S_{11} \leq -10$ dB criterion.

We may conclude that our models fail to predict the real and imaginary components of S_{11}. It shows that more complex models might be needed to handle non-linear relationships in the design space.

Model Performance and Interpretability:

- How do PCA and regression models compare in terms of their ability to simplify and predict the antenna's performance?
- What are the potential limitations of these models, and how could they be improved to more accurately represent complex, nonlinear electromagnetic behavior?

PCA aims to reduce the number of predictors in the design space by retaining most of the variance. Despite being an effective technique, PCA lies on strong assumptions and has several limitations. It assumes that relationship between predictors is linear [14]. Therefore, one limitation is that if predictors have nonlinear relationship, we obtain a wrong model by PCA [14]. Another limitation is the loss of interpretability [14]. When we project the data in the new space, it is more difficult to understand the results in the context of antenna design and performance since the transformed representations lack obvious physical meaning [14]. As stated above, no prior information of the target variable is necessary for PCA [6]. Therefore, although it may simplify our model by reducing the dimensionality of the design space, it does not guarantee that we can make a prediction about the antenna’s performance. If we could interpret the principal component as the feature index for the antenna’s performance, we would relate the design space and S_{11} in terms of prediction. However, we failed to do so. The design space seems to be randomly generated. Additionally, due to its sensitivity to outliers, PCA may produce less reliable representations by altering the principal components [14]. As an improvement, Kernel PCA can be used to handle nonlinear relationships between the predictors [15].

On the other hand, the multiple linear regression is also used to predict the antenna’s performance. Linear regression is a supervised machine learning technique [16]. It is simple to compute and implement. One significant advantage of linear regression is its interpretability [16]. The coefficients in a linear regression model directly quantify how each design parameter influences the real or imaginary part of S_{11}. An increase in the design parameter with a positive coefficient results in the increase in S_{11}, while an increase in the design parameter with a negative coefficient results in the increase in S_{11}. The main assumption in this model is that the relationship between predictor and response variables is linear [17]. Therefore, our models at selected frequencies are unable to capture nonlinearity and fail to predict the antenna's performance, especially for the imaginary part of S_{11}. It may also be possible that there are interactions between design variables (i.e. $x_1.x_2$), which is not considered in this model. As an improvement, nonlinear electromagnetic behavior can be captured by advanced models such as polynomial regression [18]. Additionally, integrating physical principles used in the antenna theory to our model to predict S_{11} parameter can better represent the underlying relationships.

References

[1] "S-Parameters for antennas (S11, S12, ...)," [Online]. Available: https://www.antenna-theory.com/definitions/sparameters.php#:~:text=S11%20represents%20how%20much%20power,antenna%20and%20nothing%20is%20radiated..

[2] R. "S-parameter measurement with the Moku:Lab Frequency Response Analyzer," 18 October 2024. [Online]. Available: https://liquidinstruments.com/application-notes/s-parameter-measurement/#:~:text=S%2Dparameters%20are%20complex%20numbers,as%20a%20function%20of%20frequency..

[3] [Online]. Available: https://product.tdk.com/system/files/dam/doc/content/emc-guidebook/en/eemc_basic_03.pdf.

[4] R. S. "Understanding VNAs - Antenna measurements," 6 August 2024. [Online]. Available: https://www.youtube.com/watch?v=15-hd_JjmYY.

[5] "Inside wireless: VSWR, |S11|, Return loss," 19 October 2020. [Online]. Available: https://www.youtube.com/watch?v=jgjaFzsYd88.

[6] GeeksforGeeks, "Principal Component Analysis(PCA)," 10 September 2024. [Online]. Available: https://www.geeksforgeeks.org/principal-component-analysis-pca/.

[7] Codecademy, "Principal component analysis," [Online]. Available: https://www.codecademy.com/article/principal-component-analysis-intro.

[8] J. Santos, "Principal Component Analysis (PCA)," 24 April 2024. [Online]. Available: https://julius.ai/articles/principal-component-analysis-pca.

[9] P. O. "Multiple Regression - interpretation (3of3)," 10 February 2018. [Online]. Available: https://www.youtube.com/watch?v=fYStutigCkE.

[10] J. Frost, "How to Interpret the F-test of Overall Significance in Regression Analysis," [Online]. Available: https://statisticsbyjim.com/regression/interpret-f-test-overall-significance-regression/.

[11] J. Frost, "How to Interpret P-values and Coefficients in Regression Analysis," [Online]. Available: https://statisticsbyjim.com/regression/interpret-coefficients-p-values-regression/.

[12] "Interpretation in Multiple Regression," [Online]. Available: https://www2.stat.duke.edu/courses/Spring00/sta242/handouts/beesIII.pdf.

[13] G. "Rsquared vs Adjusted Rsquared Difference," 2 November 2024. [Online]. Available: https://www.geeksforgeeks.org/r-squared-vs-adjusted-r-squared-difference/.

[14] D. "What is PCA and how can I use it?," 8 Februrary 2023. [Online]. Available: https://www.bigabid.com/what-is-pca-and-how-can-i-use-

it/#:~:text=Drawbacks%20of%20PCA%20(Principal%20Component%20Analysis)&text=PCA%20is%20also%20sensitive%20to,features%20lose%20their%20original%20meaning..

[15] "ML | Introduction to Kernel PCA," 14 April 2023. [Online]. Available: https://www.geeksforgeeks.org/ml-introduction-to-kernel-pca/.

[16] "Linear Regression in Machine learning," 23 October 2024. [Online]. Available: https://www.geeksforgeeks.org/ml-linear-regression/.

[17] "ML | Multiple Linear Regression using Python," 25 January 2023. [Online]. Available: https://www.geeksforgeeks.org/ml-multiple-linear-regression-using-python/.

[18] G. "Understanding Nonlinear Regression with Examples," 31 January 2024. [Online]. Available: https://www.geeksforgeeks.org/non-linear-regression-examples-ml/.

I used ChatGPT in my code, which is allowed. I asked several questions about:

- How to read the files
- How to plot
- How to apply PCA
- How to find the frequency with maximum variance
- How to find the local minima
- How to apply linear regression
- …

In case of need, I can provide my reasoning with my prompts.