Correlations of Material Properties

Alistair Mayfield, Dr. Chris Weinberger, Anik Faisal
Department of Mechanical Engineering
Colorado State University, Fort Collins, CO, USA
Introduction

Goal: Find, visualize, and interpret the correlations between properties of alloys

- Alloys are substance mixtures of different metal elements
  - The development of steel marked the industrial revolution & further alloy advancements will contribute to societal advancements
- Nearly infinite number of potential alloys, each with their own properties
  - This data focused on thermodynamic properties
- The strength of the relationship between two variables is commonly represented by a correlation coefficient.
  - There are several types of correlation coefficients, each with their own benefits and drawbacks.

Data from 80 alloys

- Previous experiments had found:
  - Temperature of phase transition
  - Amount of heat per mole used
- Other properties can be determined from either element make-up, experimental data, or both
  - Ex: Mass, number of valence electrons, entropy
Methods

Initial data considerations
- 18 properties that defined the alloys:
  - Ni, Co, Mn, Sn, Ga, In, Cu, Sb, B, Ti, Fe
  - M
  - VEC
  - Ttrans
  - Hmol & Hmass
  - Smol & Smass
- In order to prevent bias, all properties were used.

Correlation examinations
- Pearson Correlation ($r$)
  - Best suited for monotonic linear data.
- Spearman’s rank Correlation ($\rho$)
  - Best suited for any monotonic data.
- Comparing the two coefficients may indicate a linear relationship.

Building and Visualizing Correlation Matrix
- Formatting the data in order to build the correlation matrix.
- 18 columns as properties, 80 rows for each alloy.
- Using R functions in order to create and visualize the matrix for each type of correlation.

Analysis
- Examining:
  - Correlation trends
  - Possible correlation outliers
  - Differences between coefficient types
- Hypothesizing on potential causes
- Determining possibilities for further investigation.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$
Pearson Correlation (r) Matrix

6 Properties: Ttrans, Hmol, Hmass, Smol, Smass, and VEC
- All 6 properties showed strong positive correlation with each other.
- Ga, In, and Mass were weakly negatively correlated with the 6 properties.
- B, Sb, and Mn were positively correlated with the 6 properties.
- All other elements were not correlated with the 6 properties.

Mass
- In had a strong positive correlation with mass. Mn and Cu had weak negative correlation with mass.
- Ga, B, Ti all had weak negative correlation with mass.
- All other elements were not correlated with mass.

Elements
- Very few correlations between any element and In, Cu, Sb, B, Ti, and Fe
  - Exceptions of Sb with B & In with Sn and Ga.
- Ni, Co, Mn, Sn, and Ga seemed to have strong correlations with each other.
  - Mn was strongly negatively correlated with Ni and Ga, but Ni and Ga were strongly positively correlated.
The Spearman correlation matrix is nearly identical to the Pearson correlation matrix, with most correlation coefficients changing less than 0.01.

- Correlation between M and Ti had a decrease of 0.27 from r and ρ.
- The correlation between B and Sb had an increase of 0.21 from r and ρ.
Additional Results

Three metal composition correlations that stuck out to me from the correlation matrixes.
Next Steps

One potential future experiment could more data sets where all alloys have one of the 11 elements in their make up. Many of the alloys do not contain more than a few of the metals, which may have had an impact on some of the correlations.

There are also several other properties in these alloys that could be found through further experimentation, which could lead to further correlations.

Even without further experimental data, there is still a large amount of scientific literature available that could be compared to the data. This could expand our understanding of the correlations and their potential causations.

Conclusions

Overall, there are quite a few trends within the data that show a relationship between properties of the alloy and its element makeup.

Some of these have a direct causation relationship, such as the entropy being calculated from the transition temperature and heat.

A given an alloy made up of some combination of the 11 elements could have most of its properties (T_trans, H_mol, S_mol, etc) predicted, mainly by calculating of the average VEC then using the correlations of the data.
What benefits did you get from your SURE experience?

I started the project feeling a bit uncomfortable, due to it being open ended. However, I feel that I have learned quite a bit about how to work on them both individually and with other people over the course of my SURE research.

Additionally, I have learned how to apply R and MATLAB to data sets for statistical analysis.

References and Acknowledgements

This work was supported by a cooperative agreement between Colorado State University and U.S. Army Research Laboratory under contract number W911NF-19-2-0252.


Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84).

Thank you to the Suzanne and Walter Scott Foundation, Tointon Family Foundation, The Filsinger Family, Caterpillar Inc., and Contributors to the Dean’s Innovation fund for making the SURE program possible.
Thank you