Discovering Phase transitions using unsupervised machine learning PCA

Jairo J. Orozco Sandoval, Jun-Qiang Lu
University of Puerto Rico, Department of Physics, Mayaguez, PR 00681, USA

Introduction

In recent years, machine learning has been a helpful tool to discover new physics without prior human knowledge, specifically in complex problems in physics. This means we can know how a physical system behaves even without knowing its Hamiltonian. It has been possible to use machine learning to predict crystal structures, approximate density functions, model molecular atomization energy, and many other applications. For many years classifying and discovering phases and phase transitions is one of the most important topics in Condensed Matter Physics [1]; however, it is not an easy job to do, especially when we work with complex systems and the number of states is very large. One of the most important models in the theoretical physics to study phase transitions is the Ising model, given it was the first model that could successfully predict a phase transition [5]. Motivated by the search of physical properties through the machine learning techniques, in this work, we apply unsupervised machine learning technique, the Principal Component Analysis (PCA) to study square and hexagonal ferromagnetic lattice Ising systems, in order to recognize phases, phase transitions, and related physical properties without knowing any information about the microscopic theory or the order parameters. We study the Ising model in square lattice as our toy model, once it was studied, we apply PCA in hexagonal lattice and later compare the results.

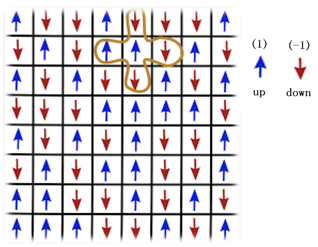
The Ising model

Square system

Proposed by Lenz in 1920

Discussed by Ernst Ising in 1925

The exact solution of the two-dimensional Ising model was made by Onsager in 1944



$$T_c = 2J/K_B \ln(\sqrt{2}+1) = 2.2692(J/K_B)$$

$$H = -J\sum_{<i,j>} S_i S_j - h\sum_i S_i$$

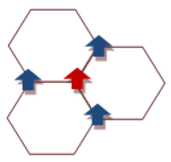
$J > 0$ *ferromagnetic*

$J < 0$ *antiferromagnetic*

Hexagonal system

The interactions are only between the three nearest neighbors

Hexagonal Lattice



$$H = -J\sum_{<i,j>} S_i S_j - h\sum_i S_i$$

$$T_c = 2J/K_B \ln(2+\sqrt{3}) = 1.519(J/K_B)$$

Finite size scaling

In the Ising model the 1st derivative of the free energy gives the magnetization M and the 2nd derivative gives the magnetic susceptibility χ_m

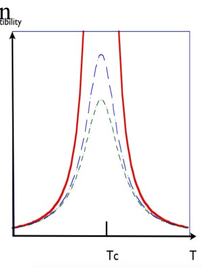
$$M = \frac{1}{V}\left(\frac{dF}{dH}\right)_{h=0} \qquad \chi_m = \frac{1}{V}\left(\frac{d^2F}{dH^2}\right)_{h=0}$$

Magnetization, $M \sim |T - T_C|^{\beta}$
Magnetic susceptibility, $\chi \sim |T - T_C|^{-\gamma}$
Heat capacity, $C \sim |T - T_C|^{-\alpha}$
Correlation length, $\xi \sim |T - T_C|^{-\nu}$
The critical exponents for the 2D Ising model are known exactly:

$\beta = 0.125$ $\alpha = 0$ $\gamma = 1.75$ $\nu = 1$

- In an infinite system ($L \sim \infty$)
- In a finite size the correlation length is $\xi \sim L$.
- The system has a pseudocritical point when

$$\left|T_{C(\infty)} - T_{C(L)}\right|^{-\nu} \sim L$$



- Then the susceptibility,

$$\chi \sim L^{\gamma/\nu}$$

Monte Carlo method

The idea of the method is to find an algorithm to generate a long sequence of configurations of a system, such that after a while each configuration is generated with the adequate probability to describe the equilibrium of the system

We start with a disorder random state from the temperature after the critical point to the equilibrium temperature.

Pick a random single spin k to be flipped.

Calculate the difference in energy before and after the flip $E_\nu - E_\mu$.

The change in energy between the two state is thus

$$E_\nu - E_\mu = -J\sum_{\langle ij\rangle} S_i^\nu S_j^\nu + J\sum_{\langle ij\rangle} S_i^\mu S_j^\mu$$

As we flip a single spin, most of the terms in the calculation in the energy different don't change and the different of energy is reduce to

$$E_\nu - E_\mu = -J\sum_{i\ n.n\ to\ k} S_i^\mu (S_k^\nu - S_k^\mu)$$

If the spin that we chose is $S_k^\mu = +1$, then after it has been flipped, we have $S_k^\nu = -1$, then $S_k^\nu - S_k^\mu = -2$ or in the other case that $S_k^\mu = -1$ and $S_k^\nu = +1$,

$$S_k^\nu - S_k^\mu = +2,$$

$$S_k^\nu - S_k^\mu = -2S_k^\mu$$

And so,

$$E_\nu - E_\mu = 2JS_k^\mu \sum_{i\ n.n\ to\ k} S_i^\mu$$

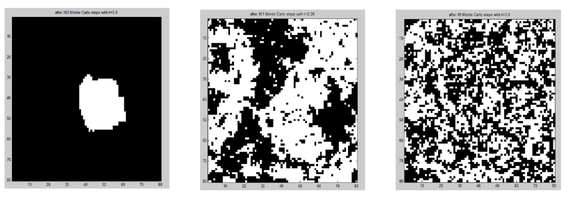
If $E_\nu - E_\mu \le 0$ we definitely accept the move and flip the

spin $S_k \to -S_k$.

If $E_\nu - E_\mu > 0$ we still want to flip the spin with probability

$$\mathrm{A}(\mu \to \nu) = e^{-\beta(E_\nu - E_\mu)}.$$

We choose a random number r between zero and one. If the number r is less that our acceptance ratio $r < \mathrm{A}(\mu \to \nu)$, then we flip the spin. If it isn't, we leave the spin alone. This process is repeated over and over. Choosing spin, calculating the energy change to see if we flip it, and then deciding whether to flip it according to the acceptance ratio.



T < T c T ≈ T c T > T c

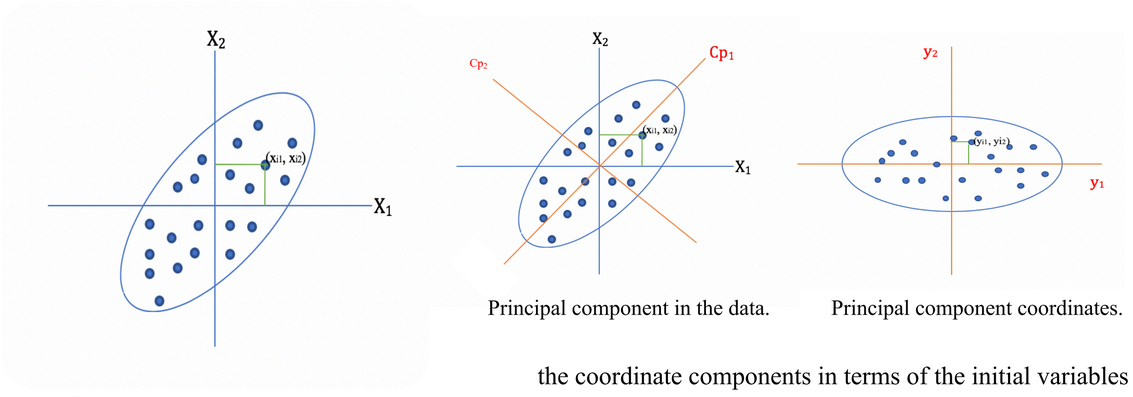
Monte Carlo simulation of ferromagnetic Ising model below Tc, over Tc, and above Tc in a finite size scale. The black dots are the spins -1 and the whites are spins +1

Lattice	Lattice size (L)	M.C steps	T_{min}/T_{max} *(T/J)*	ΔT
Square	10,20,30,40,50	30000	0.8/3.5	0.025
Hexagonal	20,40,60,80	30000	0.82/2.08	0.01

- Ferromagnetic interaction $J = 1$
- Square system = 5000 configuration per temperature (109) having a total of configuration of 545000
- Hexagonal system = 5000 configuration per temperature (124) having a total of configurations of 620000

Principal Component Analysis (PCA)

- PCA is a powerful tool to compress your data into small dimension without to much loss of information



Representation of data in the variables. Principal component in the data. Principal component coordinates.

the coordinate components in terms of the initial variables are:

$$y_{i1} = x_{i1}w_{11} + x_{i2}w_{21} \qquad y_{i2} = x_{i1}w_{12} + x_{i2}w_{22}$$

For all the coordinates component we have

$$\begin{pmatrix} y_{1j} \\ . \\ . \\ . \\ y_{1N} \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nN} \end{pmatrix} \begin{pmatrix} w_{1j} \\ . \\ . \\ . \\ w_{Nj} \end{pmatrix}$$

for all the data set of components

$$\begin{pmatrix} y_{11} & \cdots & y_{1N} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nN} \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nN} \end{pmatrix} \begin{pmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{N1} & \cdots & w_{NN} \end{pmatrix}$$

PCA in the Ising model

Data was collected by a Monte Carlo simulation in matrix X, with dimensions MxN, where $M = nT$, and N is the lattice size.

$$X = \begin{pmatrix} 1 & 1 & -1 & \cdots & -1 & 1 & -1 \\ & \vdots & & \ddots & & \vdots & \\ -1 & 1 & -1 & \cdots & 1 & 1 & -1 \end{pmatrix}_{MxN}$$

$$Y = \boldsymbol{X}\boldsymbol{W}$$

The orthogonal transformation is due to vector $W = (w_1; w_2; \ldots; w_N)$, where $w's$ are called weights; the first weight is found by:

$$w_1 = \arg max_{\|w\|=1}\left\{\sum_i (x_i . w)^2\right\}$$

the eigenvector corresponding to the largest few eigenvalues of the matrix $\boldsymbol{X^T X}(covariance)$ can be found by:

$$X^T X w_n = \lambda_n w_n$$

The principal components are calculated as:

$$Y_n = X w_n$$

where $\boldsymbol{w_1}$ will be the vector corresponding to the largest variance, namely, the largest eigenvalue.

The results are in based in the 'quantified principal components' that are defined as the average over all the configuration n at the same temperature:

$$\langle |y_n| \rangle = 1/n \sum_n |y_n|$$

Simulation details

Data from Monte Carlo Method
Square = 545000 X L^2
Hexagonal = 620000 X L^2

PCA was performed on Python

Results

PCA results of square lattices

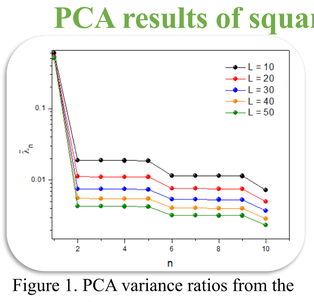


Figure 1. PCA variance ratios from the Ising configurations for square lattice

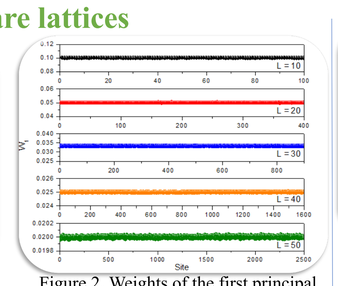


Figure 2. Weights of the first principal component for each lattice size for square lattice.

Notice that $w_1 \backsimeq 1/L$.
Then, given a configuration from one of the rows of **X**.

$$Y_1 = \frac{1}{L}[s_1, s_2, s_3 \ldots]\begin{bmatrix} 1 \\ 1 \\ 1 \\ \ldots \end{bmatrix} = \frac{1}{L}\sum_{i=1}^{N} s_i$$

If a lattice with only two spin interactions is considered and s_1 is the spin of the first lattice site

$$x_{ord} = (s_1, s_1)$$

and for the disordered phase:

$$x_{dis1} = (s_1, s_1) \qquad x_{dis2} = (s_1, -s_1)$$

When PCA is fed with configurations of both phases, the covariance to be computed is:

$$C = X^T X = \frac{1}{M}\sum_n x_n^T x_n$$

where, M is the total number of configurations and x_n is the nth configuration.
With enough data, it can be written:

$$C_{ord} = \langle x_{ord}^T x_{ord} \rangle$$

Where < > denotes the average over the value that x_{ord} can take, then for this case:

$$C_{ord} = \tfrac{1}{2}\textstyle\sum_{s_1} s_1^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

For C_{dis}:

$$C_{dis} = \frac{1}{2}\left(\langle x_{dis1}^T x_{dis1}\rangle + \langle x_{dis2}^T x_{dis2}\rangle\right) = \mathrm{I}$$

C_{ord}, yields to the eigenvalues with its corresponding eigenvectors.

$$\lambda_1 = 2, \qquad w_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 0, \qquad w_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

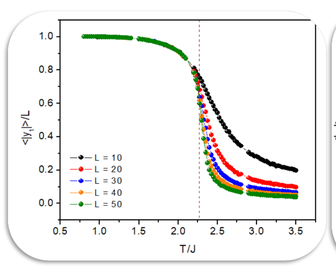


Figure 3. The quantified first leading component versus temperature which represents the magnetization of the system

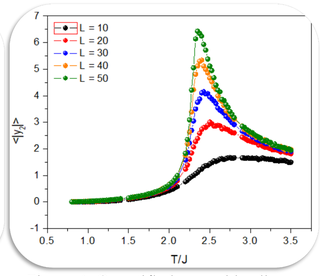


Figure 4. Quantified second leading component versus temperature which represents the susceptibility of the system

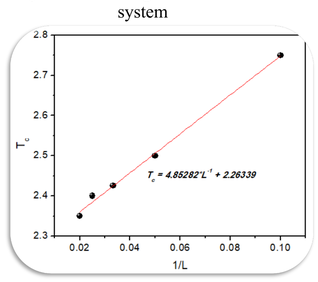


Figure 6. Critical temperatures taken from the maximums of the susceptibility versus the inverse of the lattice size.

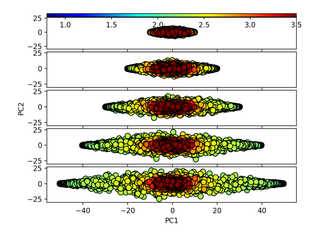
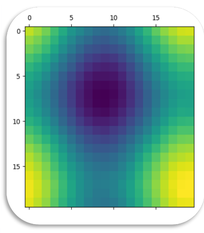
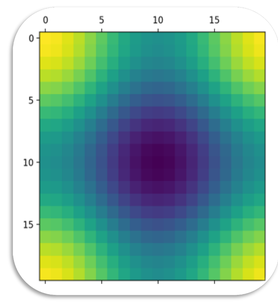


Figure 5. Projection of the spin configurations onto the plane for the two principal components for lattice of size 10,20,30,40 and 50 with 300 configurations for each temperature.



weight vector of the second component (w_2)



$$W_2' = \frac{1}{L}[\cos(r_1 k_1), \ldots, \cos(r_N k_1)] + \frac{1}{L}[\cos(r_1 k_2), \ldots, \cos(r_N k_2)]$$

where r_i is the lattice site and $k_1 = (0, 2\pi/L)$, $k_2 = (2\pi/L, 0)$ are the lowest Fourier wave vectors. The first component is associated with the origin in $k_0 = (0,0)$.

PCA results of hexagonal lattices

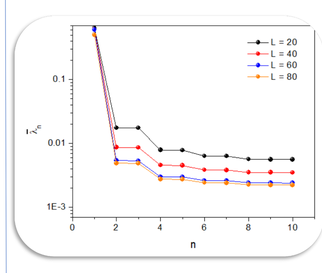


Figure 7. PCA first explained variance ratios from the Ising configurations for hexagonal lattice

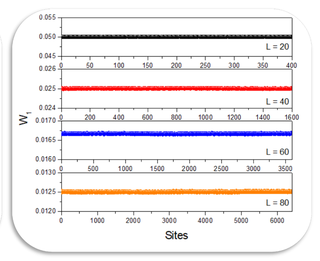


Figure 8. The Weights of the first principal component for each lattice size.

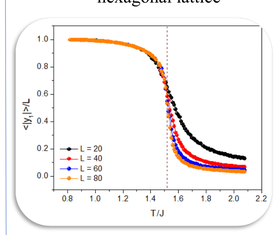


Figure 9. The quantified first leading component versus temperature for hexagonal lattice which represents the magnetization of the system

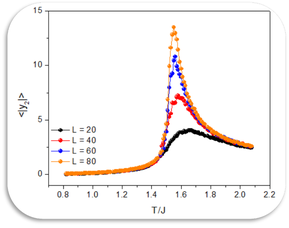


Figure 10. The quantified second leading component versus temperature for hexagonal lattice which represents the susceptibility of the system.

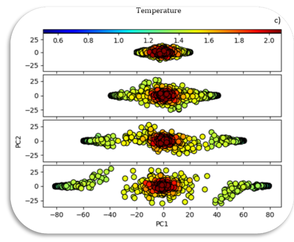


Figure 11. Projection of the spin configurations onto the plane for the two principal components for lattice of size 20,40,60,180 with 100 configurations for each temperature.

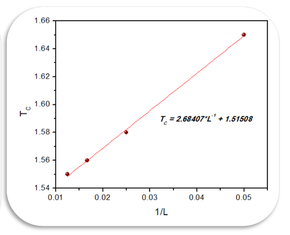
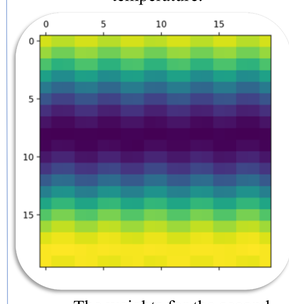
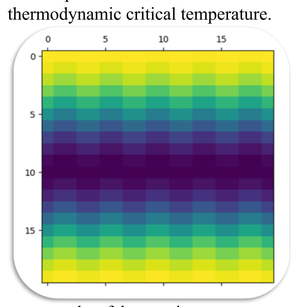


Figure 12. Critical temperature predicted in the Ising hexagonal system ($T_c \approx 1.51508\, J/K_B$). This result is close to the exact theorical value $T_c \approx 1.519\, J/K_B$, representing a 0.5% percent error from the true thermodynamic critical temperature.



The weights for the second component vector plotted on the hexagonal lattice L=20



plot of the equation:

$$W_{2hex}' = \cos\frac{1}{L}[\cos(r_1 k), \ldots, \cos(r_N k)]$$

where r_i is the hexagonal x-coordinate lattice site and $k = (2\pi/L, 0)$.

Conclusions

PCA shows a principal component and a constant weight vector, related to this particular component, for both systems. From this main component, was possible to identify the order parameter of the system and was possible to mimic the magnetization. Also, the critical temperature T_c in square and hexagonal systems were determined. For the square system a $T_c = 2.26339\, J/K_B$ was obtained and for the hexagonal system a $T_c = 1.51508\, J/K_B$, having a 0.5% percent error from the true thermodynamic critical temperature. When PCA was feed with spin configurations from Monte Carlo, I could recognize order patterns by clustering the data between the ordered and disordered phases. An interesting fact about PCA technique for the ferromagnetic Ising model is how the weight vectors correspond to the Fourier modes of the spins configuration. In the ordered phase, the physics of the Ising model is enclosed in a single point $k_0 = (0,0)$; for this reason, a single dominant eigenvalue is shown. Careful analysis reveals that further components also encase relevant information about the system. For example, information about the susceptibility of the Ising model could be obtained from the second component. The weight vector associated with this component shows that the second Fourier mode corresponding to the spin's configurations remains an unknown that can be addressed by future research. The others Fourier mode can potentially explain the physical meaning of other missing components, which could contribute to the Ising ferromagnetic model.

References

[1] L. Wang, *Discovering Phase Transitions with Unsupervised Learning.* Beijing National Lab for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China (2016).

[2] A. Rebelo, Unsupervised learning of physical models: Uses and limitations of Principal Components Analysis, master thesis, Institute of Theorical Physics, (2017).

[3] J. Carrasquilla and R. Melko, Machine learning phases of matter, Nat. Phys. 13, 431-434 (2017).

[4] W. Hu, R. Singh and R. Scalettar, Discovering Phases, Phase Transitions and Crossovers through Unsupervised Machine Learning: A critical examination, Phys. Rev. E 95, 062122 (2017).

[5] L. Onsager. *A two- dimensional model with an order-disorder transition*, Physical Review 65 117 (1944).