AI & Machine Learning in Physics

PHYS3151 (6 credits)

Time & Place : Tue 13:30-14:20, 14:30-15:20 MW T6
Fri 14:30-15:20 EH 102

Teachers: Zi Yang Meng (zymeng@hku.hk), HOC 231

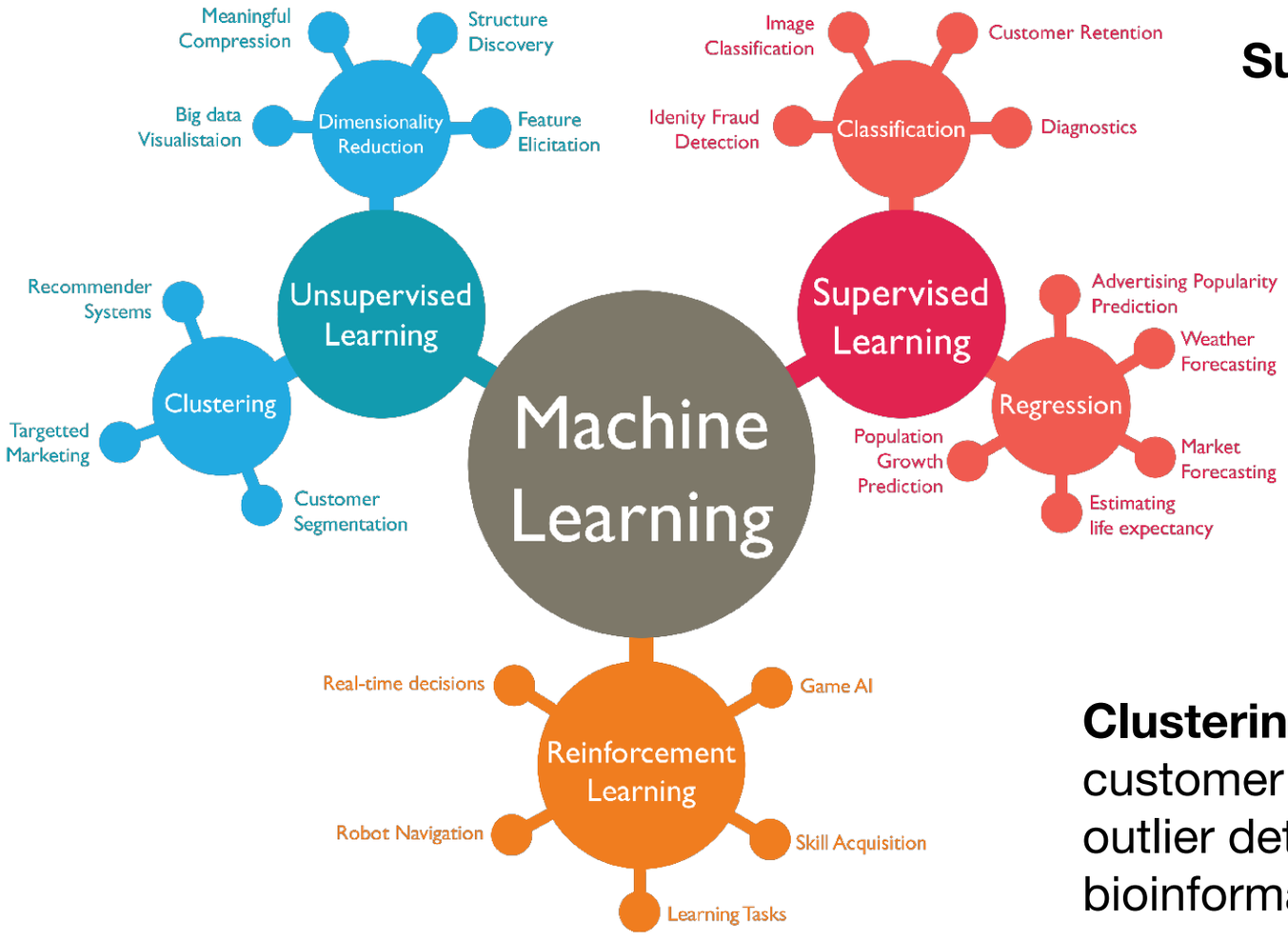
https://quantummc.xyz/hku-phys3151-machine-learning-in-physics-2024/

Tutor: Min Long (minlo@connect.hku.hk), HOC 217

Content

Content

AI & Machine Learning Basics



Supervised Learning: Classification & Regression

Input —> machine/model —> Output
Correct outputs are provided by the supervisor

Unsupervised Learning: only have input data

Find regularities from the input

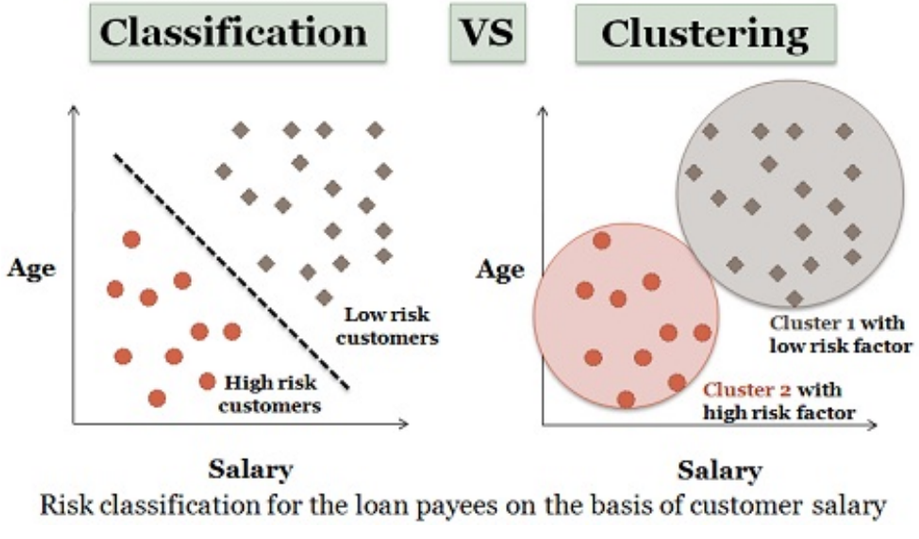
Clustering:
customer segmentation, customer relationship management, outlier detection; image compression
bioinformatics: DNA, RNA, amino acids, Motif, Proteins, sequence alignments

Good references from towards data science

[a-one-stop-shop-for-principal-component-analysis]

https://setosa.io/ev/principal-component-analysis/

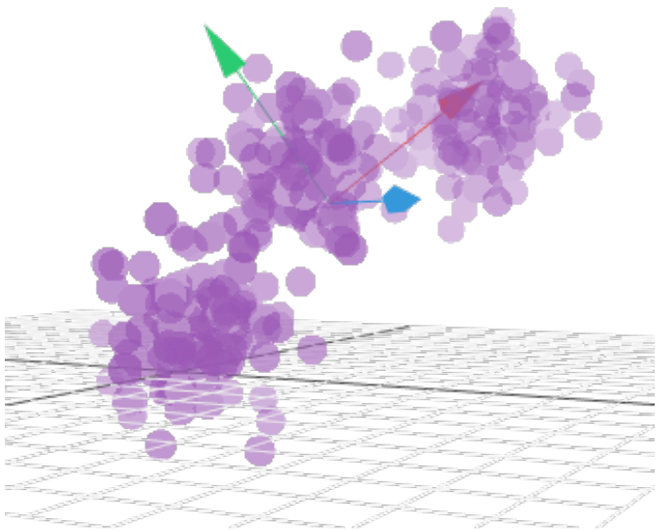
Understanding Principal Component Analysis

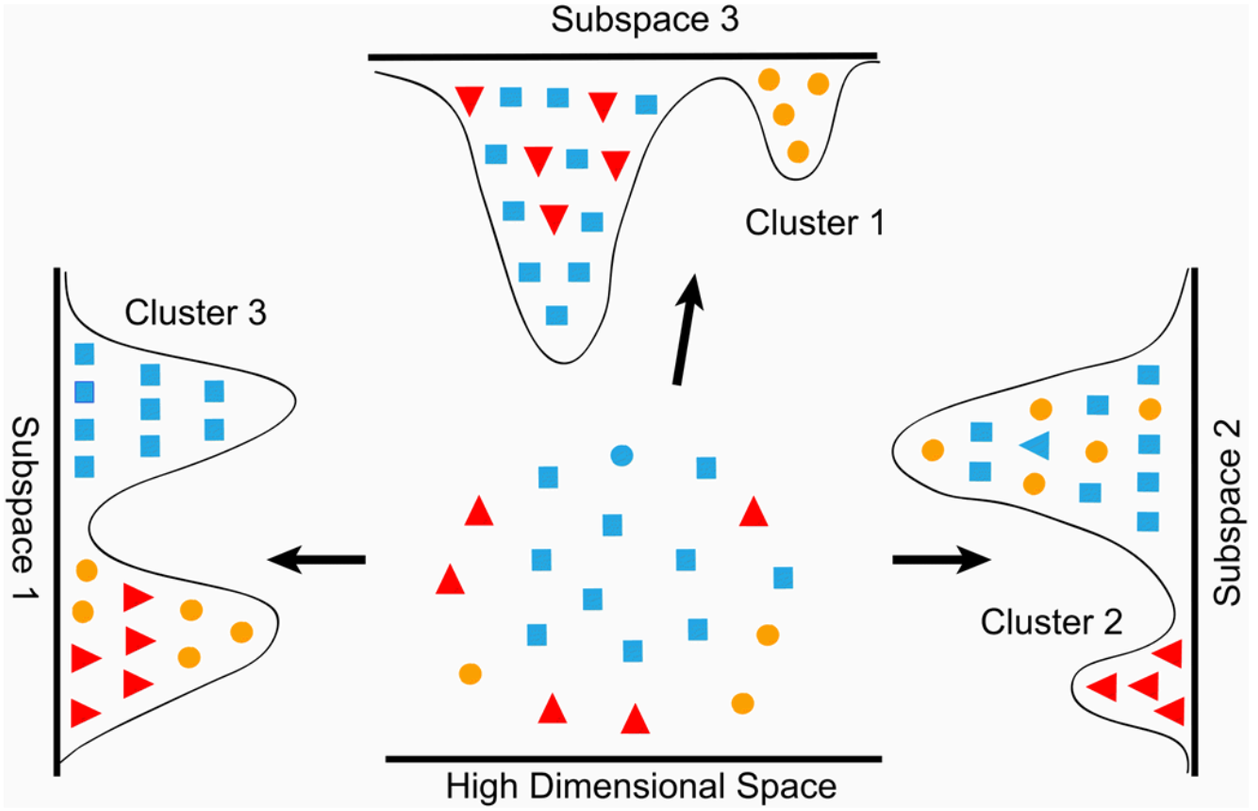


Risk classification for the loan payees on the basis of customer salary

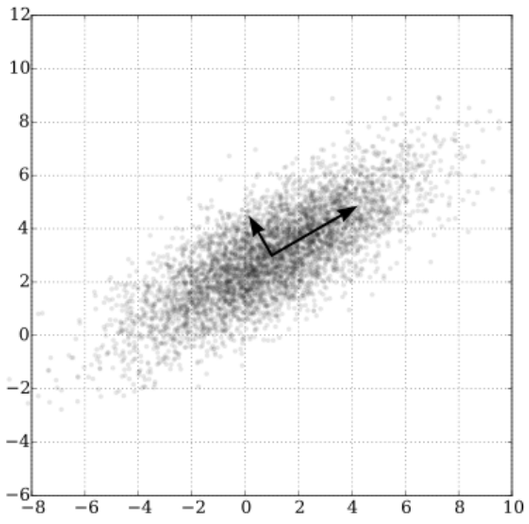
Dimension Reduction and Feature Extraction

- Reduce the dimensionality of the problem, decrease the complexity;
- Simpler models are more robust;
- Knowledge extraction





- Feature Selection/Elimination:
  - Advantage: find k out of N dimension dataset and discard the other N-k features
  - Disadvantage: gain no information from those variables dropped



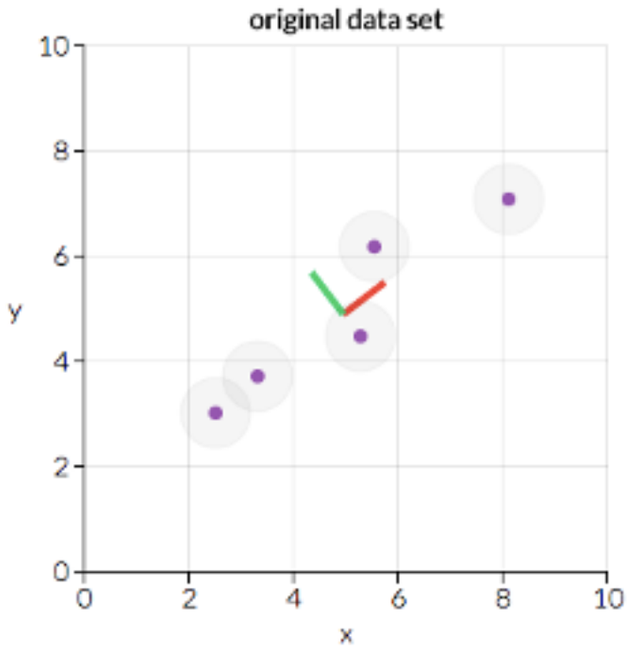
- **Feature Extraction:**
  - Find new set of k-dimensions that are combinations of the original N-dimension

- **Principal Component Analysis (PCA)** is a technique for feature extraction.
  - Advantage: each of the new variables after PCA are all independent of one another

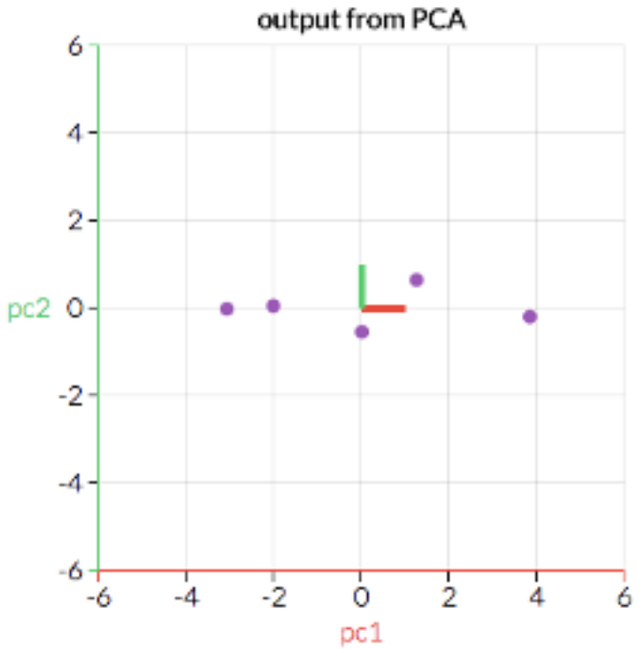
Principal Component Analysis

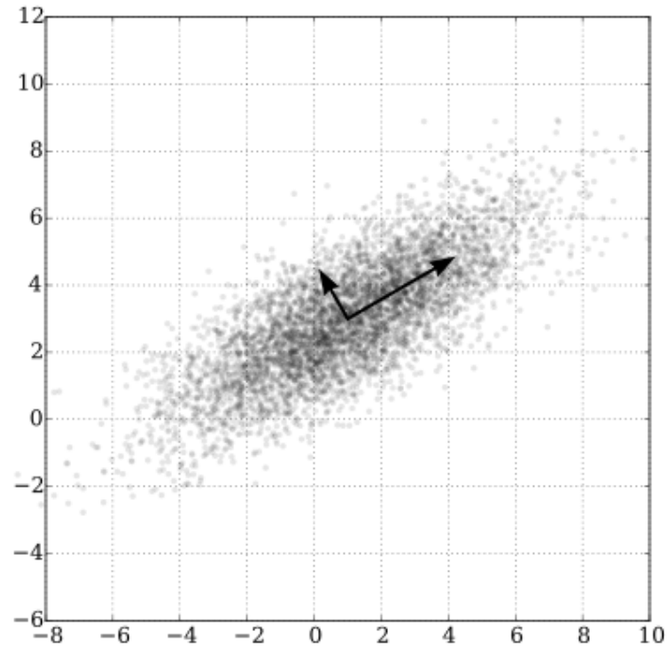
- Find linearly independent k dimensions which can represent the N dimension data
- More important = more variance / more spread out data
- Those newly found k dimensions should allow us to predict/reconstruct the original dimensions, with minimal projection/reconstruction error

Which direction is more important ?



Then, let's move to that direction





By projection, we reduce the dimensionality of feature space

Principal Component Analysis

$$X = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_N^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_N^{(2)} \\ \vdots & & \vdots & \\ x_1^{(M)} & x_2^{(M)} & \cdots & x_n^{(M)} \end{bmatrix}$$

Mean value $\mu_j = \frac{1}{M}\sum_{i=1}^{M} x_j^{(i)}$

Variance $\sigma_j^2 = \frac{\sum_{i=1}^{M}(x_j^{(i)} - \mu_j)^2}{M}$

$j = 1, 2, \cdots, N$

$$Z = \begin{bmatrix} \frac{x_1^{(1)}-\mu_1}{\sigma_1} & \frac{x_2^{(1)}-\mu_2}{\sigma_2} & \cdots & \frac{x_N^{(1)}-\mu_N}{\sigma_N} \\ \frac{x_1^{(2)}-\mu_1}{\sigma_1} & \frac{x_2^{(2)}-\mu_2}{\sigma_2} & \cdots & \frac{x_N^{(2)}-\mu_N}{\sigma_N} \\ \vdots & & \vdots & \\ \frac{x_1^{(M)}-\mu_1}{\sigma_1} & \frac{x_2^{(M)}-\mu_2}{\sigma_2} & \cdots & \frac{x_n^{(M)}-\mu_N}{\sigma_N} \end{bmatrix}$$

Covariance matrix

$$\Sigma = \frac{1}{M} Z^T Z = \begin{bmatrix} \frac{x_1^{(1)}-\mu_1}{\sigma_1} & \frac{x_1^{(2)}-\mu_1}{\sigma_1} & \cdots & \frac{x_1^{(M)}-\mu_1}{\sigma_1} \\ \frac{x_2^{(1)}-\mu_2}{\sigma_2} & \frac{x_2^{(2)}-\mu_2}{\sigma_2} & \cdots & \frac{x_2^{(M)}-\mu_2}{\sigma_2} \\ \vdots & & \vdots & \\ \frac{x_N^{(1)}-\mu_N}{\sigma_N} & \frac{x_N^{(2)}-\mu_N}{\sigma_N} & \cdots & \frac{x_N^{(M)}-\mu_N}{\sigma_N} \end{bmatrix} \times \begin{bmatrix} \frac{x_1^{(1)}-\mu_1}{\sigma_1} & \frac{x_2^{(1)}-\mu_2}{\sigma_2} & \cdots & \frac{x_N^{(1)}-\mu_N}{\sigma_N} \\ \frac{x_1^{(2)}-\mu_1}{\sigma_1} & \frac{x_2^{(2)}-\mu_2}{\sigma_2} & \cdots & \frac{x_N^{(2)}-\mu_N}{\sigma_N} \\ \vdots & & \vdots & \\ \frac{x_1^{(M)}-\mu_1}{\sigma_1} & \frac{x_2^{(M)}-\mu_2}{\sigma_2} & \cdots & \frac{x_n^{(M)}-\mu_N}{\sigma_N} \end{bmatrix} /M$$

$$= \begin{bmatrix} \sum_{i=1}^{M}(\frac{x_1^{(i)}-\mu_1}{\sigma_1})^2 & \sum_{i=1}^{M}\frac{x_1^{(i)}-\mu_1}{\sigma_1}\frac{x_2^{(i)}-\mu_2}{\sigma_2} & \cdots & \sum_{i=1}^{M}\frac{x_1^{(i)}-\mu_1}{\sigma_1}\frac{x_N^{(i)}-\mu_N}{\sigma_N} \\ \sum_{i=1}^{M}\frac{x_2^{(i)}-\mu_2}{\sigma_2}\frac{x_1^{(i)}-\mu_1}{\sigma_1} & \sum_{i=1}^{M}(\frac{x_2^{(i)}-\mu_2}{\sigma_2})^2 & \cdots & \sum_{i=1}^{M}\frac{x_2^{(i)}-\mu_2}{\sigma_2}\frac{x_N^{(i)}-\mu_N}{\sigma_N} \\ \vdots & & \vdots & \\ \sum_{i=1}^{M}\frac{x_N^{(i)}-\mu_N}{\sigma_N}\frac{x_1^{(i)}-\mu_1}{\sigma_1} & \sum_{i=1}^{M}\frac{x_N^{(i)}-\mu_N}{\sigma_N}\frac{x_2^{(i)}-\mu_2}{\sigma_2} & \cdots & \sum_{i=1}^{M}(\frac{x_N^{(i)}-\mu_N}{\sigma_N})^2 \end{bmatrix} /M$$

$\mathbb{R}^{N \times N}$

Symmetric
positive semidefinite matrix

$$\sum = cov[x_i, x_j] = E[(x_i - E[x_i])(x_j - E[x_j])] \qquad j = 1, 2, \cdots, N$$

Principal Component Analysis

- Understand the eigenvalue decomposition $\Sigma = \mathbf{PDP}^T = \mathbf{PDP}^{-1}$

$$\underbrace{\boldsymbol{\Sigma}}_{(N\times N)} = \underbrace{\begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_N \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}}_{\mathbf{P}:(N\times N)} \underbrace{\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & \cdots & \lambda_N \end{pmatrix}}_{\mathbf{D}} \underbrace{\begin{pmatrix} \leftarrow & \mathbf{u}_1 & \rightarrow \\ \leftarrow & \mathbf{u}_2 & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{u}_N & \rightarrow \end{pmatrix}}_{\mathbf{P}^T:(N\times N)}$$

P matrix of eigenvector, eigenvectors are orthogonal and linear independent

D matrix of eigenvalues

- Sort eigenvalues from large to small, sort the eigenvector accordingly $\mathbf{P}^*$
- Rotate data to the principle components $\mathbf{Z}^* = \mathbf{Z} \times \mathbf{P}^*$
- Determine how many features to keep
- That's it, PCA.

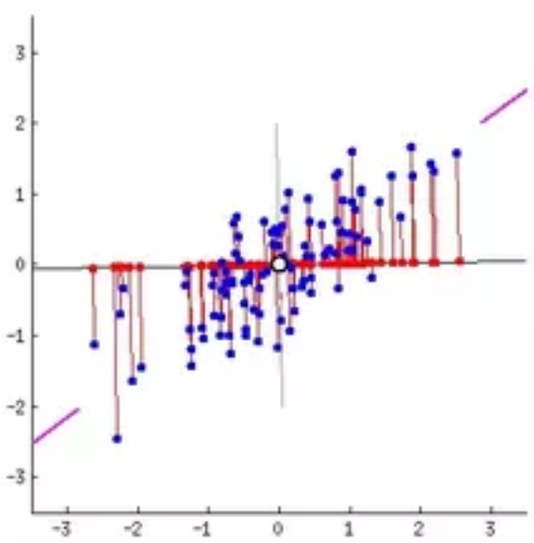
A cat, just to get your attention. :)

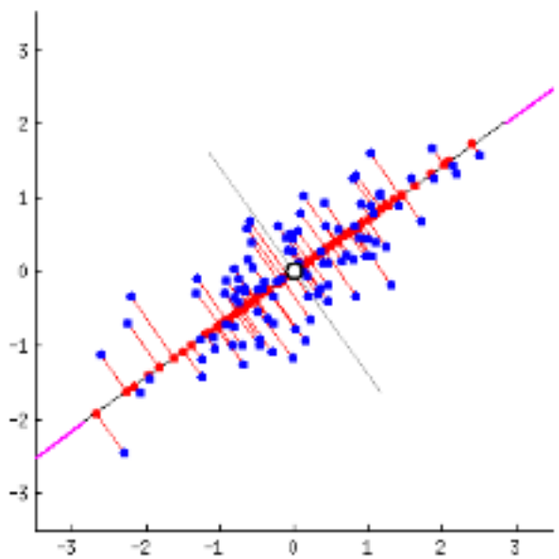
“In fact, every principal component will ALWAYS be orthogonal (a.k.a. official math term for perpendicular) **to every other principal component.** (Don’t believe me? Try to break the applet!)"

$$\mathbf{Z}^* = \mathbf{Z} \times \mathbf{P}^*$$

$$\mathbb{R}^{M \times k} \qquad \mathbb{R}^{M \times N} \qquad \mathbb{R}^{N \times k}$$

- Find new set of k-dimensions that are combinations of the original N-dimension,
- Dimension reduction





Principal Component Analysis

Thus, PCA is a method that brings together:

1. A measure of how each variable is associated with one another. (Covariance matrix.)

2. The directions in which our data are dispersed. (Eigenvectors.)

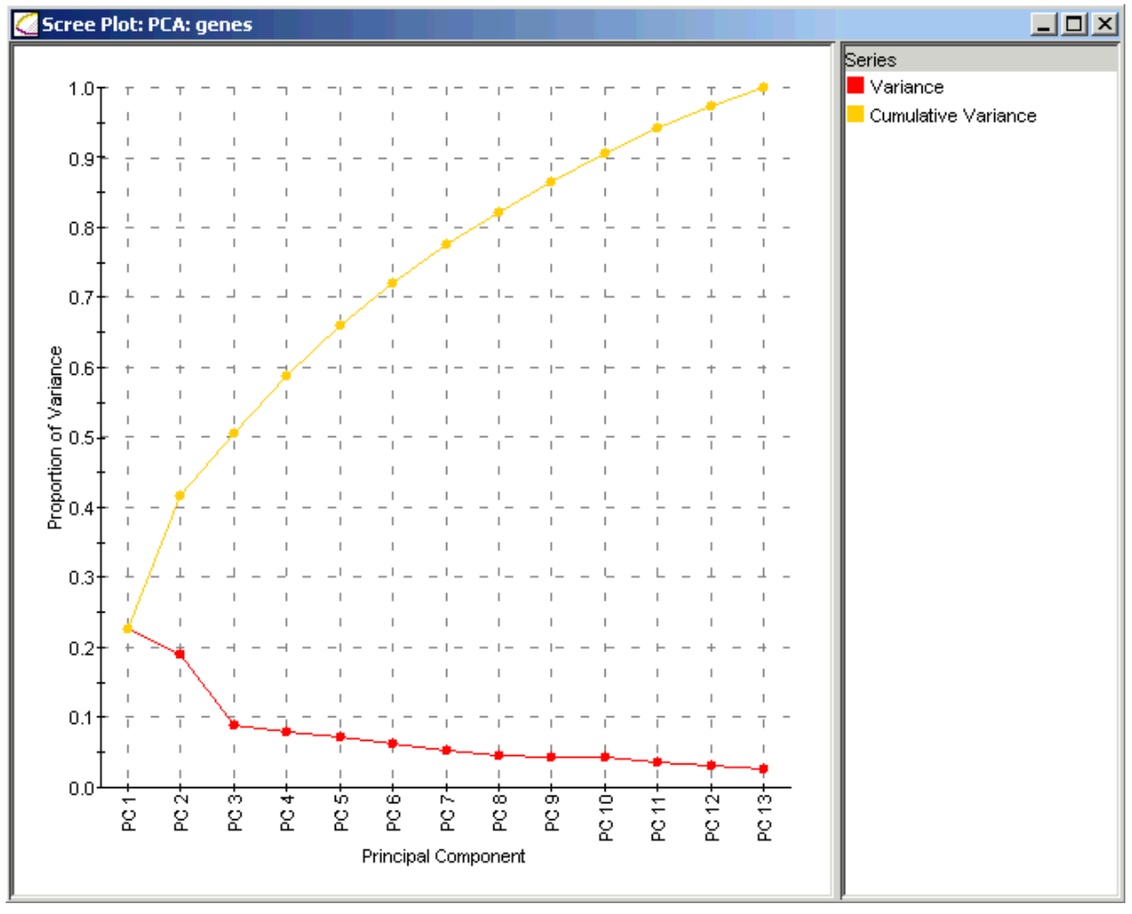
3. The relative importance of these different directions. (Eigenvalues.)

how many features to keep versus how many to drop

1. select by hand

2. calculate the proportion of variance explained

$$\text{proportion of variance} = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{N} \lambda_i}$$



npj | Quantum Materials

www.nature.com/npjquantmats

ARTICLE OPEN

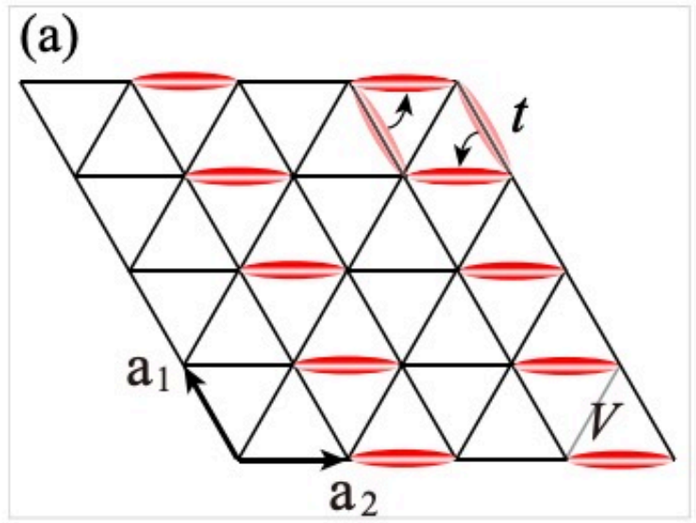
Check for updates

Topological phase transition and single/multi anyon dynamics of Z_2 spin liquid

Zheng Yan[1], Yan-Cheng Wang[2], Nvsen Ma[3], Yang Qi[4,5,6]✉ and Zi Yang Meng[1]✉

Among the quantum many-body models that host anyon excitation and topological orders, quantum dimer models (QDM) provide a suitable playground for studying the relation between single-anyon and multi-anyon continuum spectra. However, as the prototypical correlated system with local constraints, the generic solution of QDM at different lattice geometry and parameter regimes is still missing due to the lack of controlled methodologies. Here we obtain, via sweeping cluster quantum Monte Carlo algorithm, the excitation spectra in different phases of the triangular lattice QDM. Our results reveal the single vison excitations inside the Z_2 quantum spin liquid (QSL) and its condensation towards the $\sqrt{12}\times\sqrt{12}$ valence bond solid (VBS), and demonstrate the translational symmetry fractionalization and emergent O(4) symmetry at the QSL-VBS transition. We find the single vison excitations, whose convolution qualitatively reproduces the dimer spectra, are not free but subject to interaction effects throughout the transition. The nature of the VBS with its O(4) order parameters are unearthed in full scope. Our approach opens the avenue for generic solution of the static and dynamic properties of QDMs and has relevance towards the realization and detection of fractional excitations in programmable quantum simulators.

npj Quantum Materials (2021)6:39; https://doi.org/10.1038/s41535-021-00338-1



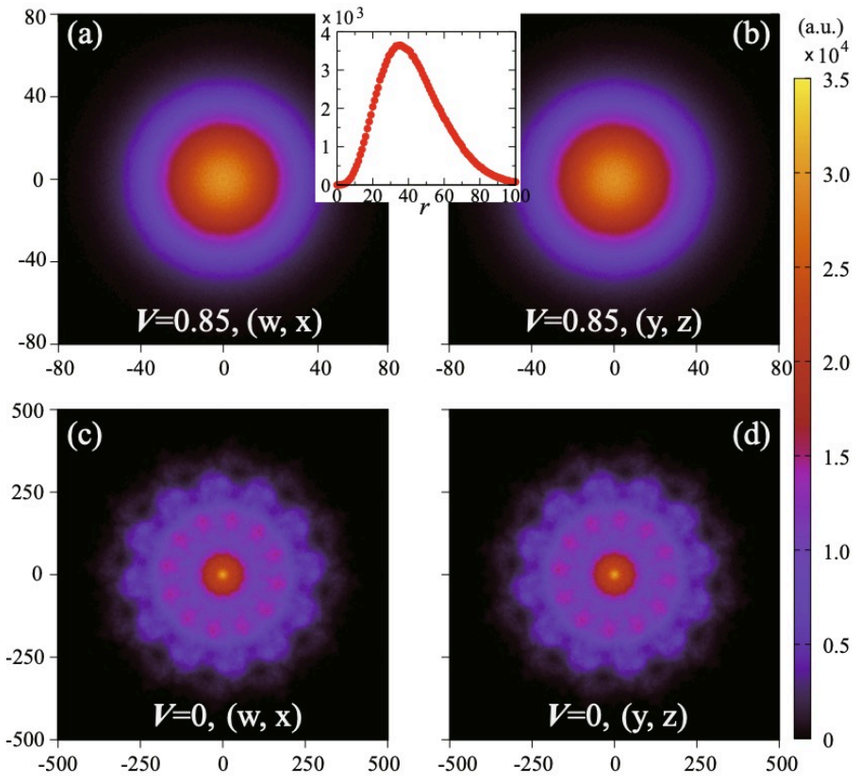


Table 1. Principal component analysis.

V	0.5	0.6	0.7	0.8	0.9	1
L_1/L_2	85.03	82.05	76.18	68.48	51.36	29.71

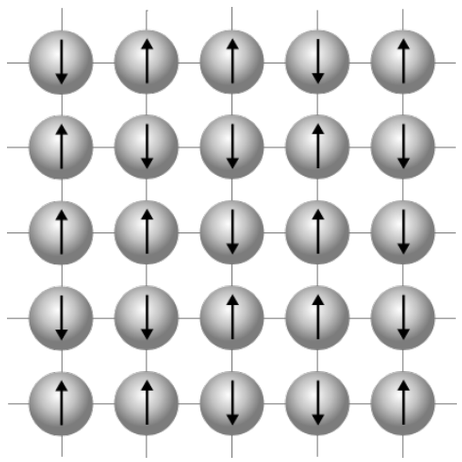
L_1/L_2 means the first largest eigenvalue over the second of the momentum-space vison correlation function matrix at B point. All data are obtained at a 12×12 lattice with $\beta = 200(T = 1/200)$.

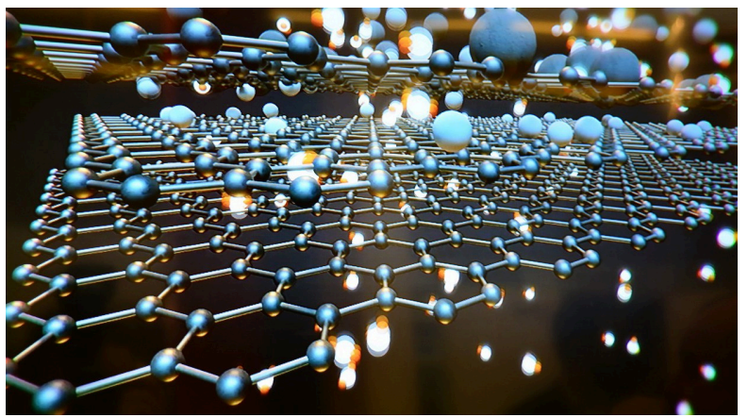
In order to numerically confirm that the order parameter ϕ_i indeed captures the QSL-VBS transition, we perform a principal component analysis (PCA) on the vison correlation function C_v to extract the condensing mode near the transition. PCA diagonalizes the 4×4 matrix of the momentum-space vison correlation function at the B point, and identifies the eigenvectors with the largest eigenvalues corresponds to the modes represented by the order parameter ϕ_i. We list the ratio of the first largest eigenvalue over the second at $V = 0.5$–1 in Table 1. Since the largest eigenvalue always dominate, it shows that the principal component of the VBS structure is indeed the expected $\sqrt{12}\times\sqrt{12}$ order. The theoretical analysis further predicts that, at the QSL-VBS critical point, the transition point acquires an emergent O(4) symmetry, as O(4)-symmetry-breaking terms become irrelevant.

AI & Machine Learning Basics

Ising model, continuous phase transition
workhorse for statistical physics

https://mattbierbaum.github.io/ising.js/

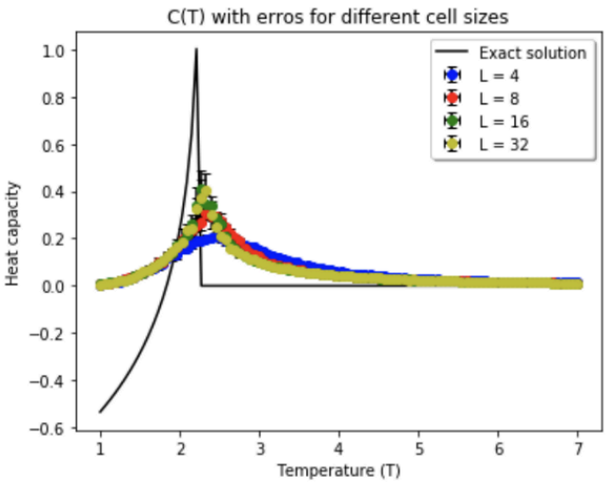


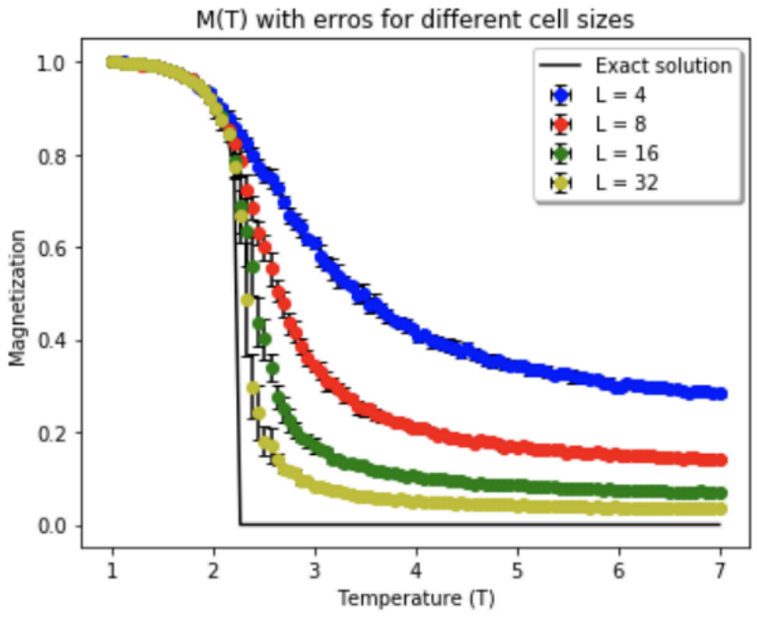


$$H = -J \sum_{\langle i,j \rangle} S_i^z S_j^z \qquad S_i^z = \pm 1$$

Configuration space: 2^N

https://en.wikipedia.org/wiki/Ising_model#/media/File:Ising_quench_b10.gif





$$m = \frac{1}{N} |\sum_{i=1}^{N} S_i^z|$$

$$m(T) = |T - T_c|^{\beta} \text{ with } \beta = 1/8 \text{ in 2D}$$