

# Singular Value Decomposition

---

**Aiichiro Nakano**

*Collaboratory for Advanced Computing & Simulations  
Department of Computer Science  
Department of Physics & Astronomy  
Department of Quantitative & Computational Biology  
University of Southern California*

**Email: anakano@usc.edu**

**Goal: Another matrix decomposition (SVD)  
for low-rank matrix approximation**

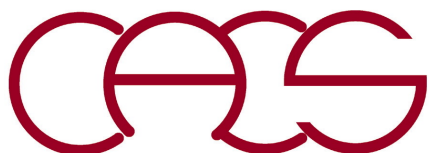
cf. Eigen decomposition

$$A = Q[\text{diagonal}]Q^T$$

QR decomposition

$$A = Q[\text{upper triangular}]$$

See note on [“least square fit”](#) & [Numerical Recipes Sec. 2.6](#)



# Rank of a Matrix

- $N \times M$  matrix  $A$  as a mapping:  $R^M \rightarrow R^N$

$$M \quad \begin{matrix} 1 \\ \left[ \begin{matrix} x \end{matrix} \right] \end{matrix} \quad x \left( \in R^M \right) \xrightarrow{A} b = Ax \left( \in R^N \right) \quad \begin{matrix} 1 \\ \left[ \begin{matrix} b \end{matrix} \right] \end{matrix} \quad N$$

- Range** of  $A$ : Vector space  $\{b = Ax | \forall x\}$
- Rank** of  $A$ : Number,  $m$ , of linearly-independent vectors in the range, *i.e.*, how many linearly-independent  $N$ -element vectors are there in the range, such that

$$b = A \forall x = \sum_{v=1}^m c_v |v\rangle$$

# Low Rank Approximations of a Matrix

- Rank-1 approximation:  $NM \rightarrow N + M$

$$\begin{array}{c}
 \begin{matrix} & & & N \times M \\ & & \overbrace{\hspace{1.5cm}} & \\ & & N \times 1 & 1 \times M \end{matrix} \\
 \begin{matrix} N \\ \psi \end{matrix} \begin{matrix} M \\ \end{matrix} \cong \begin{matrix} \begin{bmatrix} u \end{bmatrix} \begin{bmatrix} v \end{bmatrix} \end{matrix}
 \end{array}
 \quad |u\rangle\langle v| \forall x \rangle \propto |u\rangle$$

- Rank-2 approximation:  $NM \rightarrow 2(N + M)$

$$\begin{bmatrix} \psi \end{bmatrix} \cong \begin{bmatrix} u_1 \end{bmatrix} w_1 \begin{bmatrix} v_1 \end{bmatrix} + \begin{bmatrix} u_2 \end{bmatrix} w_2 \begin{bmatrix} v_2 \end{bmatrix}$$

- Rank- $m$  ( $m \ll N, M$ ) approximation:  $NM \rightarrow m(N + M)$

$$\begin{bmatrix} \psi \end{bmatrix} \cong \sum_{v=1}^m \begin{bmatrix} u_v \end{bmatrix} w_v \begin{bmatrix} v_v \end{bmatrix}$$

# Singular Value Decomposition

- **Problem:** Optimal approximation of an  $N \times M$  matrix  $\psi$  of rank- $m$  ( $m \ll N$ )?
- **Theorem:** An  $N \times M$  matrix  $\psi$  (assume  $N \geq M$ ) can be decomposed as

$$\psi = UDV^T = \sum_{v=1}^M U_{iv} d_v V_{jv} = \sum_{v=1}^M u_i^{(v)} d_v v_j^{(v)}$$

where  $U \in R^{N \times M}$  &  $V \in R^{M \times M}$  are column orthogonal &  $D$  is diagonal

$$U^T U = V^T V = I_M$$

$$\begin{matrix} & M \\ N & \left[ \begin{array}{c} \psi \end{array} \right] = \left[ \begin{array}{c} U \end{array} \right] \left[ \begin{array}{c} d_1 \\ \vdots \\ d_M \end{array} \right] \left[ \begin{array}{c} V^T \end{array} \right] \\ & N \times M \quad M \times M \quad M \times M \end{matrix}$$

See [appendix on polar & singular decompositions](#)

- **Theorem:** Sort the SVD diagonal elements in descending order,  $d_1 \geq d_2 \geq \dots \geq d_M \geq 0$ , & retain the first  $m$  terms

$$\psi^{(m)} \equiv \sum_{v=1}^m u^{(v)} d_v v^{(v)T}$$

which is optimal among  $\forall$  rank- $m$  matrices in the 2-norm sense with the error

$$\min_{\text{rank}(A)=m} \|A - \psi\|_2 = \|\psi^{(m)} - \psi\|_2 = d_{m+1}$$

cf. [singular.c](#) & [svdcmp.c](#)    cc -o singular singular.c svdcmp.c -lm

*Use the program!*

# SVD for Image Compression



Original Image



5 Iterations



10 Iterations

**D. Richards & A. Abrahamsen**



20 Iterations

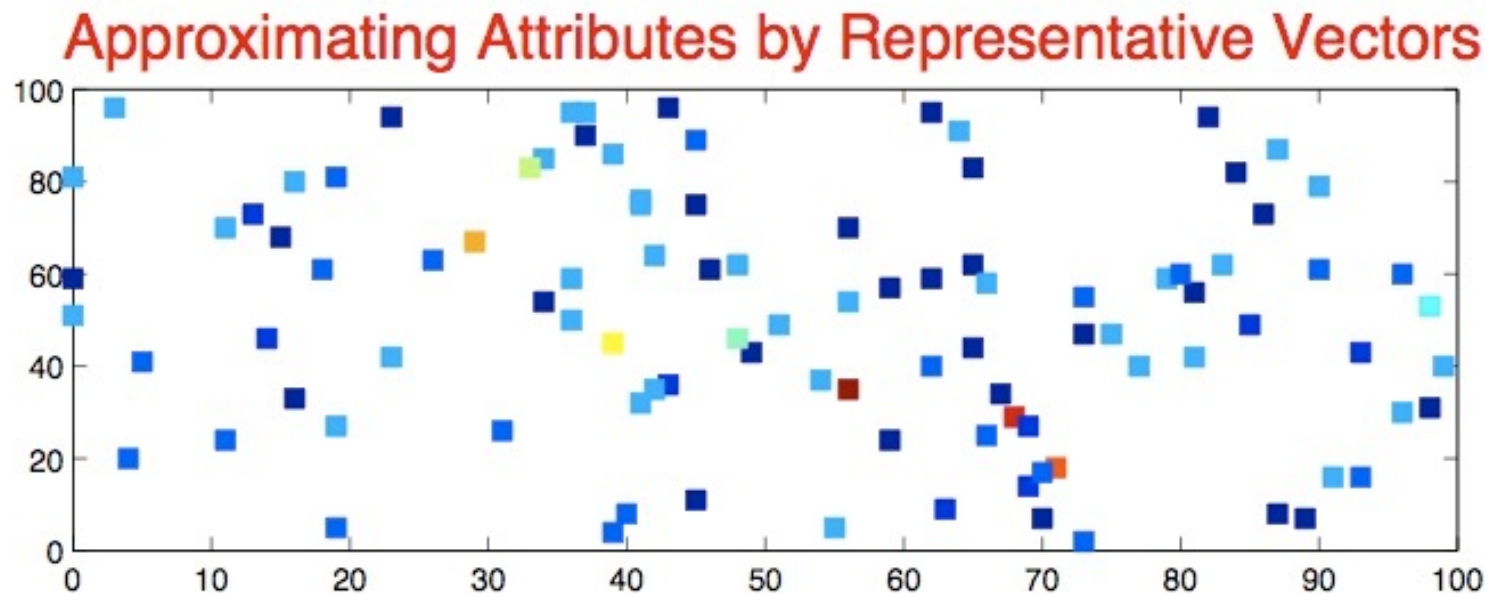
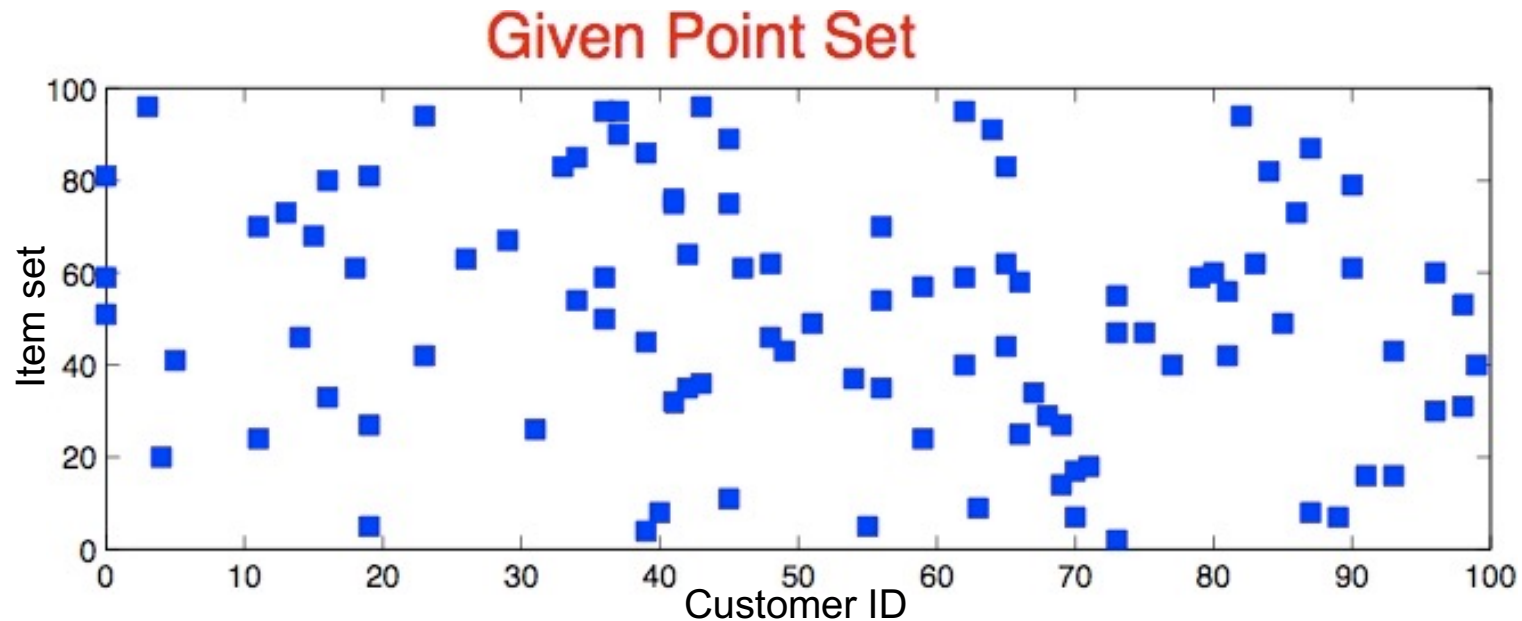


60 Iterations



100 Iterations

# SVD in Data Mining



# Reduced Density Matrix

- Quantum system coupled to an environment



$$\{|i\rangle = \psi_i(x) | i = 1, \dots, N\} \quad \{|j\rangle = \phi_j(X) | j = 1, \dots, M\}$$

- $\forall$  Quantum state of block + environment

$$|\psi\rangle = \sum_{i=1}^N \sum_{j=1}^M \psi_{ij} |i\rangle |j\rangle \quad \text{or} \quad \Psi(x, X) = \sum_{i=1}^N \sum_{j=1}^M \psi_{ij} \psi_i(x) \phi_j(X)$$

- Reduced density matrix

Arbitrary operator in the block

$$\begin{aligned} \langle \forall A \rangle &= \sum_i \sum_j \psi_{ij}^* \langle j | \langle i | A \sum_{i'} \sum_{j'} \psi_{i'j'} |i'\rangle |j'\rangle \\ &= \sum_i \sum_j \sum_{i'} \sum_{j'} \psi_{i'j'} \psi_{ij}^* \langle i | A | i' \rangle \langle j | j' \rangle \delta_{jj'} \\ &= \sum_i \sum_{i'} \sum_j \psi_{i'j} \psi_{ij}^* \langle i | A | i' \rangle \equiv \sum_i \sum_{i'} \rho_{i'i} A_{ii'} = \text{tr}_B(\rho A) \end{aligned}$$

$$\rho_{i'i} \equiv \sum_j \psi_{i'j} \psi_{ij}^* \quad A_{ii'} \equiv \langle i | A | i' \rangle$$

# Low-Rank Approx. to Reduced Density Matrix

$$\begin{aligned}\psi &\equiv \psi^{(m)} = \sum_{\nu=1}^m u^{(\nu)} d_{\nu} v^{(\nu)T} & \psi_{ij}^{(m)} &= \sum_{\nu=1}^m u_i^{(\nu)} d_{\nu} v_j^{(\nu)} \\ \rho = \psi \psi^T &\equiv \psi^{(m)} \psi^{(m)T} = \sum_{\nu=1}^m \sum_{\nu'=1}^m u^{(\nu)} d_{\nu} \left( v^{(\nu)T} v^{(\nu')} \right) d_{\nu'} u^{(\nu')T} \\ &= \sum_{\nu=1}^m \sum_{\nu'=1}^m u^{(\nu)} d_{\nu} (\delta_{\nu\nu'}) d_{\nu'} u^{(\nu')T} = \sum_{\nu=1}^m u^{(\nu)} d_{\nu}^2 u^{(\nu)T} \equiv \rho^{(m)} \\ \rho_{ii'}^{(m)} &= \sum_{\nu=1}^m u_i^{(\nu)} d_{\nu}^2 u_{i'}^{(\nu)}\end{aligned}$$

- **Density matrix renormalization group** = systematic procedure to accurately obtain a quantum ground state:
  1. Incrementally add environment to a block
  2. Solve the global (= block + environment) ground state
  3. Construct a low-rank approx. to represent the block with reduced d.o.f.

S. R. White, *Phys. Rev. B* **48**, 10345 ('93);

G. K.-L. Chan & S. Sharma, *Annu. Rev. Phys. Chem.* **62**, 465 ('11) (Eqs. 3-10)

**More SVD in quantum chemistry**

J. F. Gonthier & M. Head-Gordon, *J. Chem. Phys.* **147**, 144110 ('17)



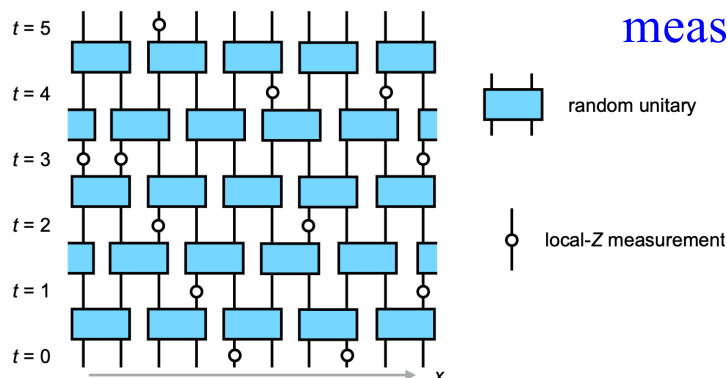
# Entanglement Entropy & Quantum Computing

- **Entanglement entropy:** A measure of quantum entanglement between two subsystems. If a state describing two subsystems A and B is a *separable* state  $|\Psi_{AB}\rangle = |\phi_A\rangle|\phi_B\rangle$  (i.e., tensor product), then the reduced density matrix  $\rho_A = \text{Tr}_B |\Psi_{AB}\rangle\langle\Psi_{AB}| = |\phi_A\rangle\langle\phi_A|$  is a *pure state*. Thus, the entropy of the state is zero. A reduced density matrix having a non-zero entropy is therefore a signal of the existence of entanglement in the system.
- **Area law:** A quantum state satisfies an *area law* if the leading term of the entanglement entropy grows at most proportionally with the *boundary* between the two partitions. Area laws are remarkably common for ground states of local *gapped* quantum many-body systems. *It greatly reduces the complexity of quantum many-body systems. The density matrix renormalization group and matrix product states, for example, implicitly rely on such area laws.*

[https://en.wikipedia.org/wiki/Entropy\\_of\\_entanglement](https://en.wikipedia.org/wiki/Entropy_of_entanglement)

## Measurement-driven entanglement transition in hybrid quantum circuits

“With increasing measurement rate, the volume law phase is unstable to a disentangled area law phase, passing through a single entanglement transition at a critical rate of measurement.”



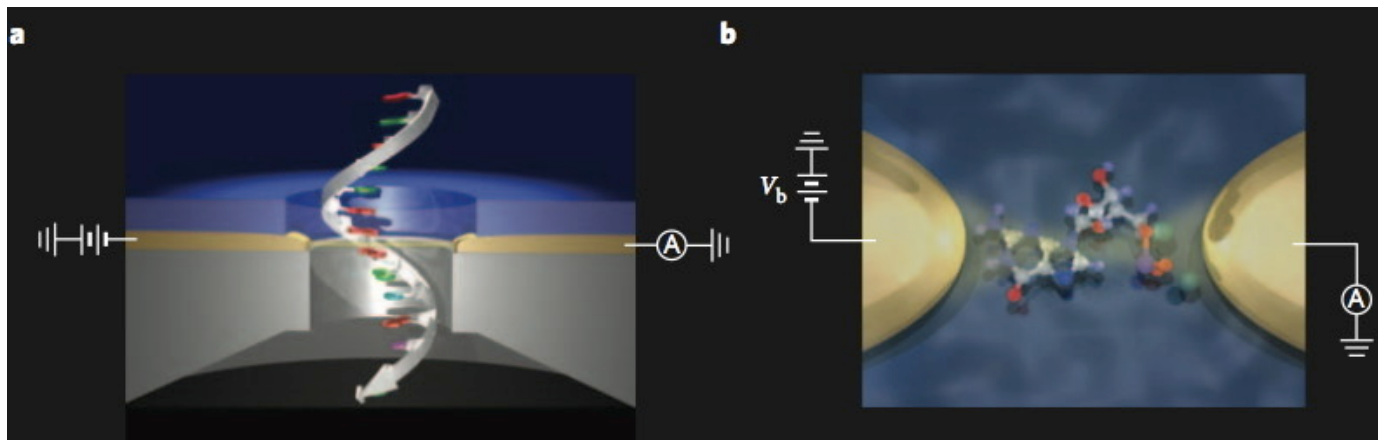
Y. Li et al., Phys. Rev. B **100**, 134306 ('19)

# SVD for Rapid Genome Sequencing

- \$10M Archon X prize for decoding 100 human genomes in 10 days & \$10K per genome (<http://genomics.xprize.org>): Preemptive attack on diseases

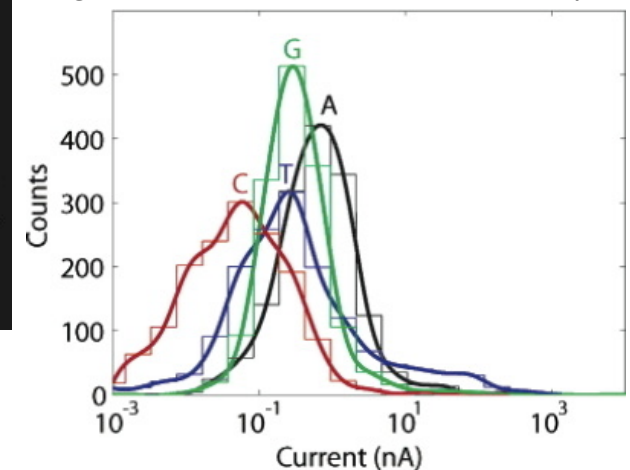


- Quantum tunneling current for rapid DNA sequencing?



Tsutsui *et al.*, *Nature Nanotechnology* ('10)

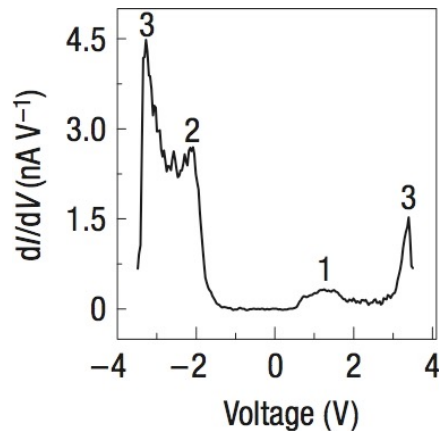
Lagerqvist *et al.*, *Nano Letters* ('06)



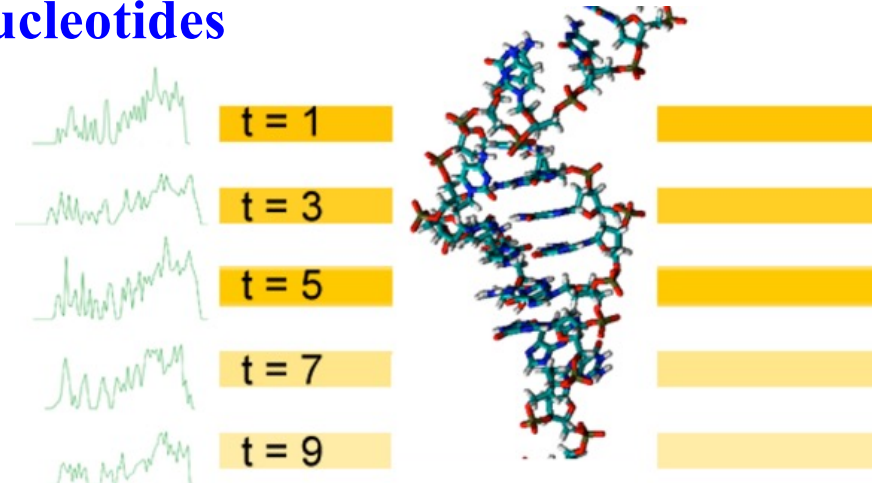
- Tunneling current alone cannot distinguish the 4 nucleotides (A, C, G, T)

# Rapid DNA Sequencing *via* Data Mining

- Use tunneling current (I)-voltage (V) characteristic (or electronic density-of-states) as the 'fingerprints' of the 4 nucleotides

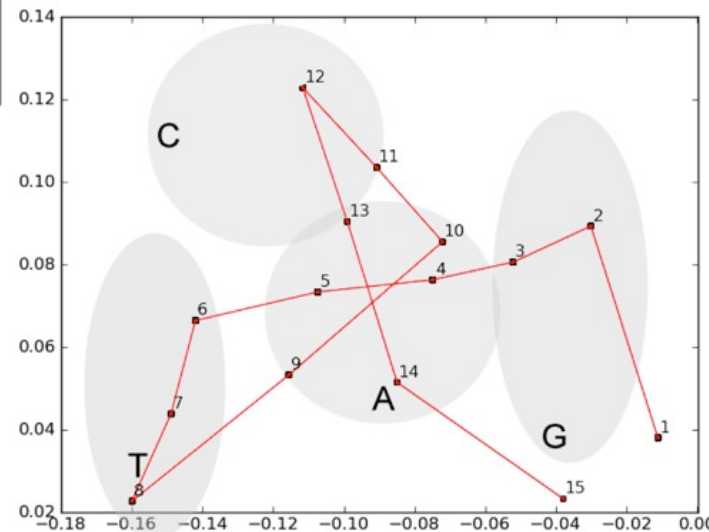
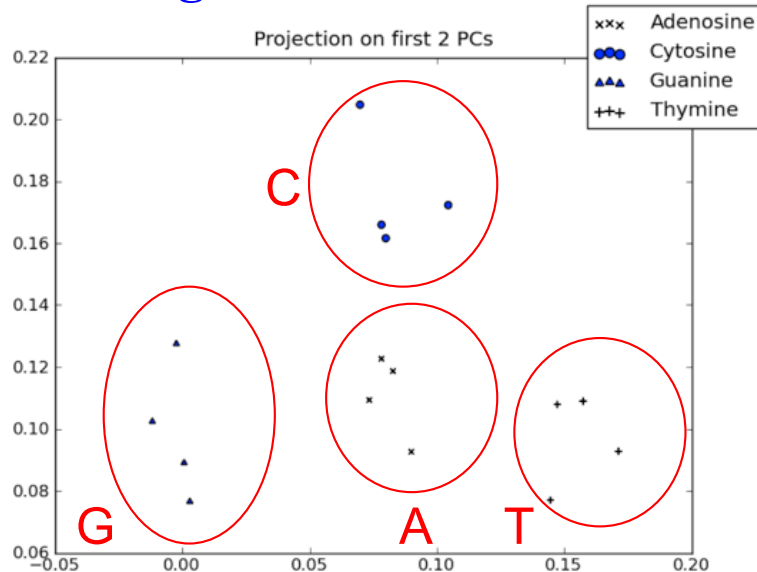


Shapir *et al.*,  
*Nature Materials* ('08)



- Principal component analysis (PCA) & fuzzy c-means clustering clearly distinguish the 4 nucleotides

H. Yuen *et al.*, *IJCS* 4, 352 ('10)



<http://www.henryyuen.net/>

- Viterbi algorithm for even higher-accuracy sequencing See [Henry's landmark discovery](#)

# SVD vs. PCA (in Economics)

- SVD of  $N$  (number of companies)  $\times$   $T$  (number of time points) of stock-price time series

$$\Xi_{T \times N}^T = \mathbf{U}_{T \times N} \Sigma_{N \times N} \mathbf{V}_{N \times N}^T$$

- Stock correlation matrix

$$\mathbf{C}_{N \times N} = \Xi_{N \times T} \Xi_{T \times N}^T$$

*Apply it in your area!*

- Principal component analysis (PCA): Eigen decomposition of the correlation matrix

$$\begin{aligned} \mathbf{C} &= \Xi \Xi^T \\ &= \mathbf{V} \Sigma \underbrace{\mathbf{U}^T \mathbf{U}}_{\mathbf{I}} \Sigma \mathbf{V}^T \\ &= \mathbf{V} \Sigma^2 \mathbf{V}^T \end{aligned}$$

- Compare the spectrum with that of random matrix theory (RMT) for judging statistical significance

