

Singular Value Decomposition

Aiichiro Nakano

*Collaboratory for Advanced Computing & Simulations
Dept. of Computer Science, Dept. of Physics & Astronomy,
Dept. of Chemical Engineering & Materials Science,
Dept. of Quantitative & Computational Biology
University of Southern California*

Email: anakano@usc.edu

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

See note on “least square fit” & Numerical Recipes Sec. 2.6



Rank of a Matrix

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

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

- **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

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

Low Rank Approximations of a Matrix

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

$$N \begin{bmatrix} M \\ \psi \end{bmatrix} \cong \begin{bmatrix} u \\ v \end{bmatrix}$$

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

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

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

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

Singular Value Decomposition

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

$$\psi = UDV^T = \sum_{\nu=1}^M U_{i\nu} d_\nu V_{j\nu} = \sum_{\nu=1}^M u_i^{(\nu)} d_\nu v_j^{(\nu)}$$

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

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

$$N \begin{bmatrix} \psi \end{bmatrix}_{N \times M} = M \begin{bmatrix} U \\ d_1 & \ddots & d_M \\ V^T \end{bmatrix}_{M \times M} \quad \text{See note on polar & singular decompositions}$$

- **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

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

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

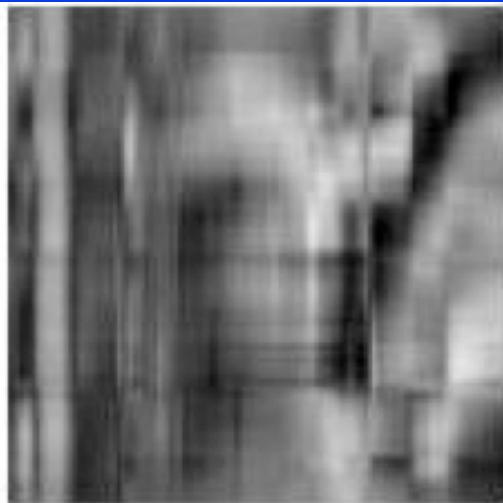
$$\min_{\text{rank}(A)=m} \|A - \psi\|_2 = \|\psi^{(m)} - \psi\|_2 = d_{m+1}$$

cf. [singular.c](#) & [svdcmp.c](#)

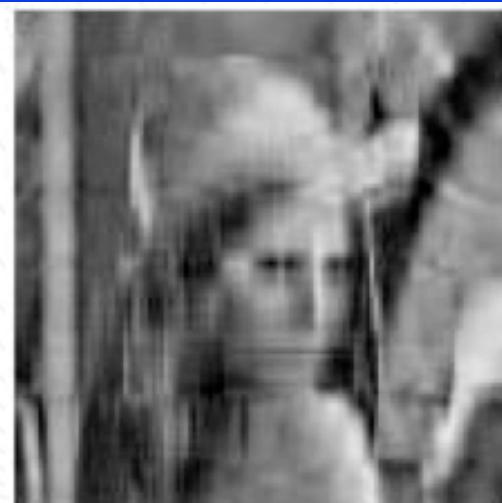
SVD for Image Compression



Original Image



5 Iterations



10 Iterations

D. Richards & A. Abrahamsen



20 Iterations

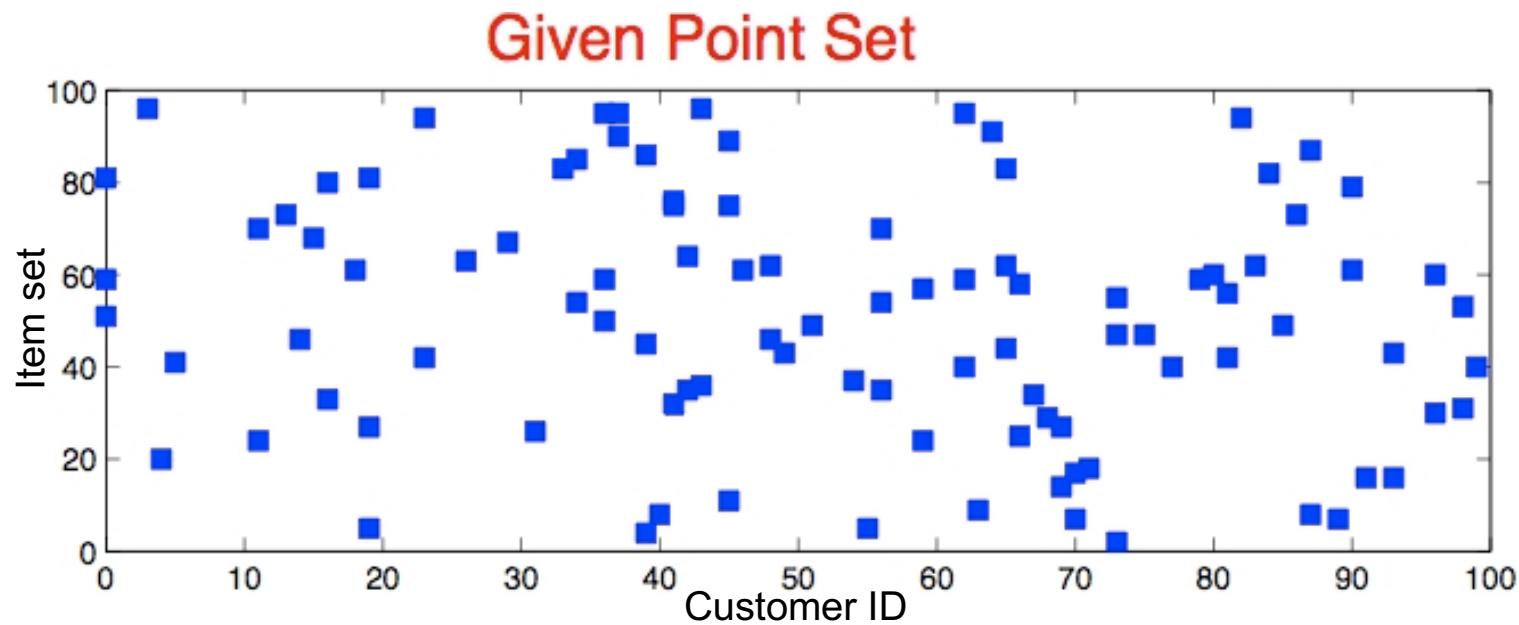


60 Iterations

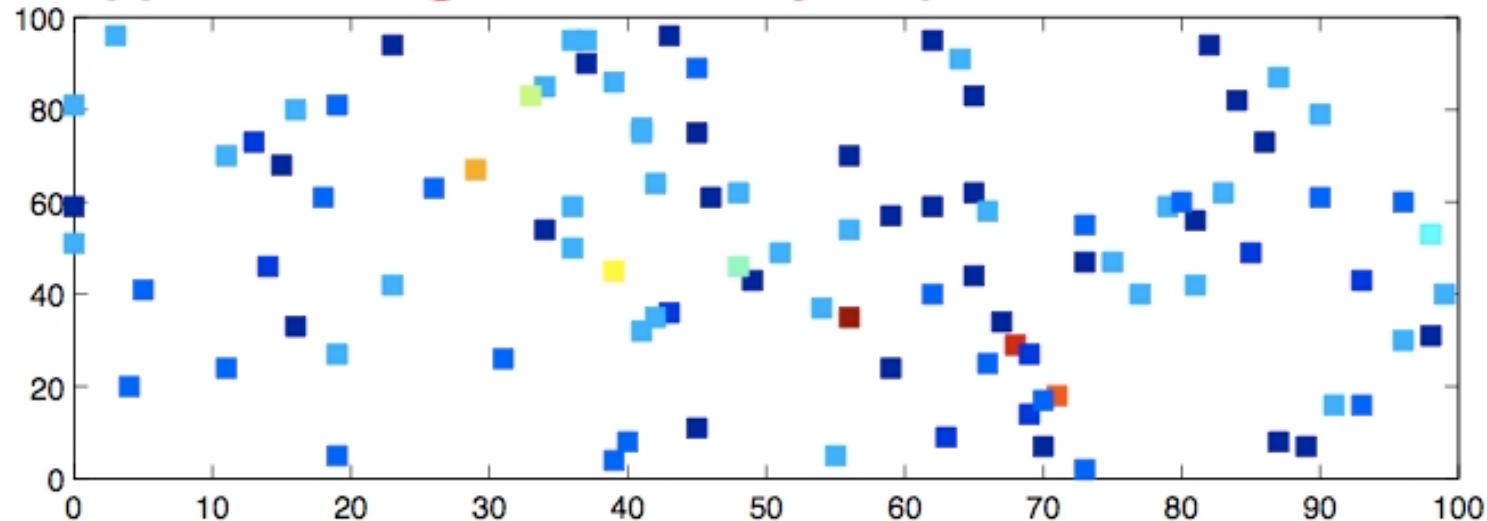


100 Iterations

SVD in Data Mining



Approximating Attributes by Representative Vectors



Reduced Density Matrix

- Quantum system coupled to an environment



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

- 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

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

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

Low-Rank Approx. to Reduced Density Matrix

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

- **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)

Entanglement Entropy

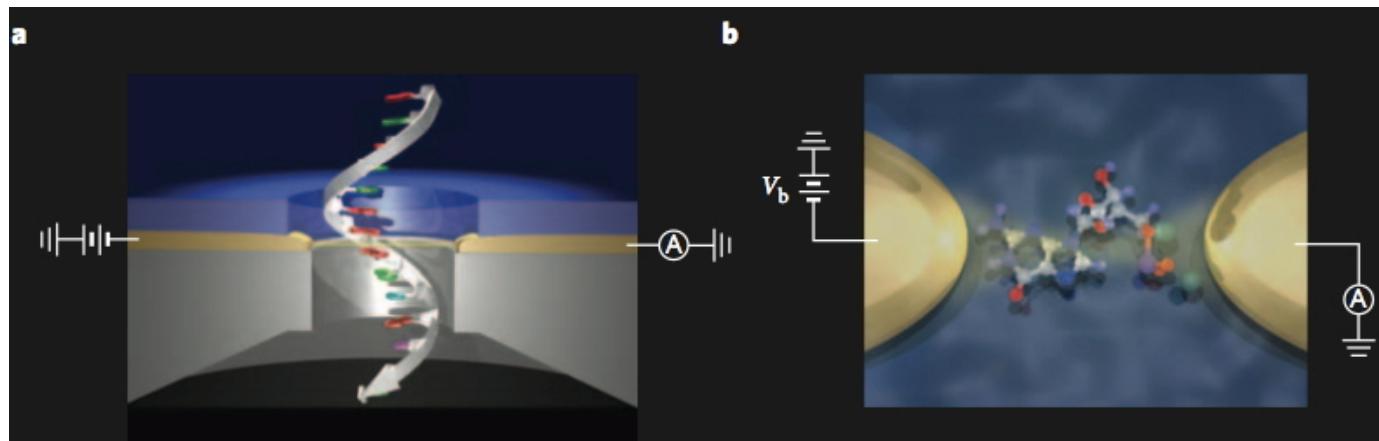
- **Entanglement entropy:** A measure of the degree of quantum entanglement between two subsystems. If a state describing two subsystems A and B is a *separable* state $|\Psi_{AB}\rangle = |\phi_A\rangle|\phi_B\rangle$, then the reduced density matrix $\rho_A = \text{Tr}_B |\Psi_{AB}\rangle\langle\Psi_{AB}| = |\phi_A\rangle\langle\phi_A|$ 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.*

Rapid Genome Sequencing

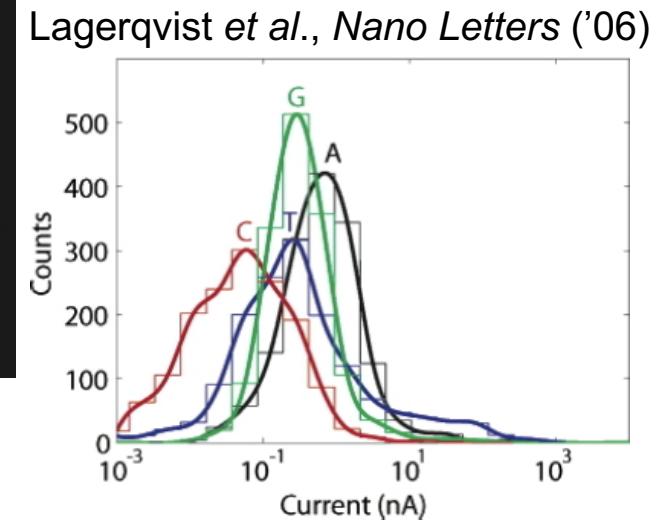
- \$10M Archon X prize for decoding 100 human genomes in 10 days & \$10K per genome (<http://genomics.xprize.org>): Preemptive attack on diseases



- Quantum tunneling current for rapid DNA sequencing?



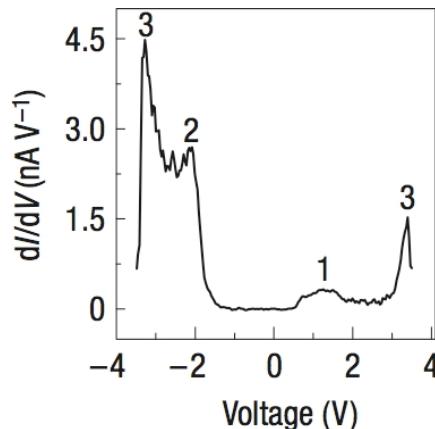
Tsutsui et al., *Nature Nanotechnology* ('10)



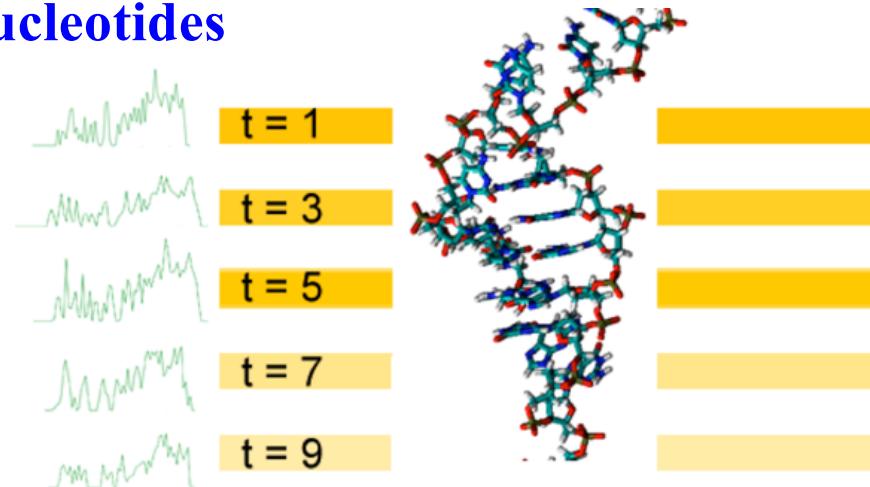
- Tunneling current alone cannot distinguish the 4 nucleotides (A, C, G, T)

Rapid DNA Sequencing via Data Mining

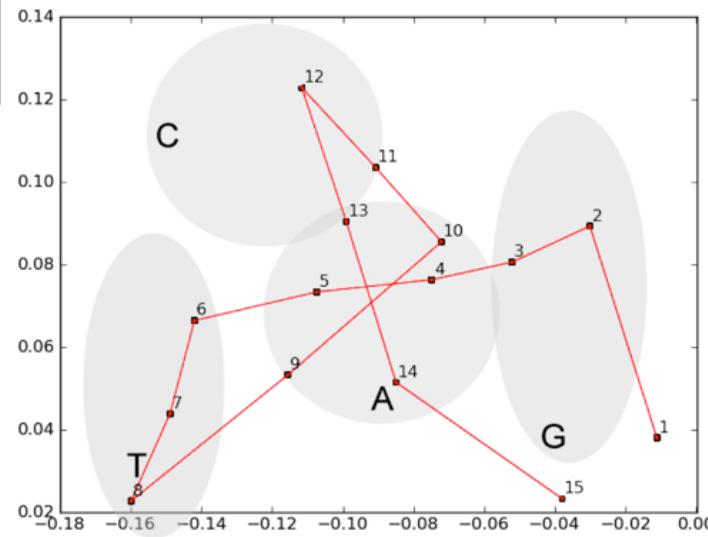
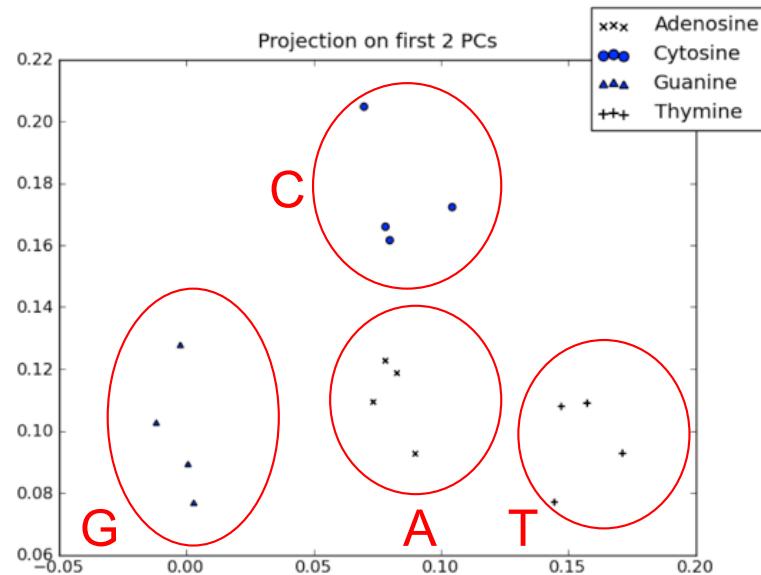
- Use tunneling current (I)-voltage (V) characteristic (or electronic density-of-states) as the ‘fingerprints’ of the 4 nucleotides



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

SVD vs. PCA

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

$$\underset{T \times N}{\Sigma}^T = \underset{T \times N}{U} \underset{N \times N}{\Sigma} \underset{N \times N}{V}^T$$

- Stock correlation matrix

$$\underset{N \times N}{C} = \underset{N \times T}{\Sigma} \underset{T \times N}{\Sigma}^T$$

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

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

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

