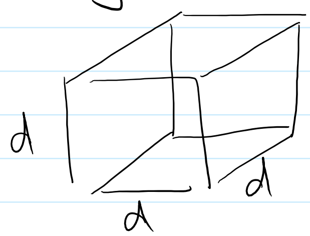
Lecture 9 Tensor decomposition

Thursday, September 22, 2016 9:52 PM

- Working with tensors
  - General Idea: convert tensors into matrices
  - slicing



take $M_{i,j} = T_{1,i,j}$

($or\ T_{i,1,j}\ T_{i,j,1}$)

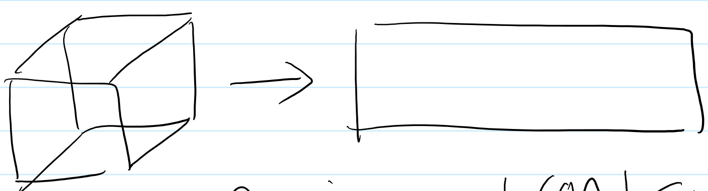
$rank(M) \le rank(T)$

Often we will need to compare more than one slices.

(recall: tensor decomposition = simultaneous matrix decomp)

Lose information (bad for sample complexity)

  - Flatening



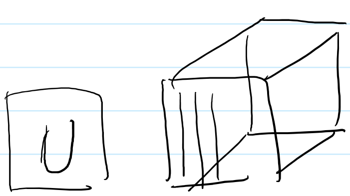
$T_{1,(2,3)}$

$M_{i,(j,k)} = T_{i,j,k}$

again, $rank(M) \le rank(T)$

$\|M\| \ge \|T\|$

  - Linear transform

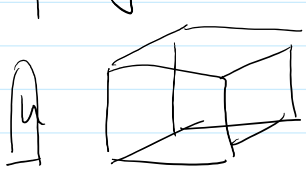


$T(U, I, I)$: apply U to the first dimension

$T(:, i, j) \leftarrow U^T T(:, i, j)$

(why U^T in the notation? See next operation)

  - Projection



$T(u, I, I)$ is a matrix M

$M_{i,j} = \langle u, T(:, i, j) \rangle$

$$= u^T T(:, i, j)$$

Special case of linear transformation, more general than slices.

- The low rank decomposition.

$$T = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i$$

- low rank form interacts well with operations above

$$T_{1,(2,3)} = \sum_{i=1}^{r} u_i \, (\underbrace{v_i \otimes w_i}_{d^2 \text{ dimensional vector}})^T$$

$$T(U, I, I) = \sum_{i=1}^{r} (U^T u_i) \otimes v_i \otimes w_i$$

$$T(u, I, I) = \sum_{i=1}^{r} \langle u, u_i \rangle v_i w_i^T \qquad (\text{matrix})$$

$$T(u, v, I) = \sum_{i=1}^{r} \langle u, u_i \rangle \langle v, v_i \rangle w_i \qquad (\text{vector})$$

$$T(u, v, w) = \sum_{i=1}^{r} \langle u, u_i \rangle \langle v, v_i \rangle \langle w, w_i \rangle \qquad (\text{number})$$

- Finding tensor decomposition

Jenrich's algorithm

Suppose $T = \sum_{i=1}^{r} u_i \otimes v_i \otimes w_i$

pick two random vectors a, b

$$M_a = T(a, I, I) = \sum_{i=1}^{r} \langle a, u_i \rangle v_i w_i^T$$

$$M_b = T(b, I, I) = \sum_{i=1}^{r} \langle b, u_i \rangle v_i w_i^T$$

v_i's = eigenvectors of $M_a M_b^{-1}$

w_i's = eigenvectors of $(M_a^{-1} M_b)^T$

pair (v_i, w_i) if their eigenvalues are reciprocals

solve for u_i (system of linear equations)

Theorem: If $\{v_i\}$ $\{w_i\}$ are linearly independent, $\{u_i\}$ are pairwise independent (corresponds to different directions), then tensor decomposition is unique. With probability 1, Jenrich's algorithm finds correct u_i, v_i, w_i (up to permutation and scaling)

- Proof: Let $V = \begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix}$ $W = \begin{bmatrix} w_1 & w_2 & \cdots & w_r \end{bmatrix}$

$$M_a = V D_a W^T \quad (D_a(i,i) = \langle a, u_i \rangle)$$

$$M_b = V D_b W^T \quad (D_b(i,i) = \langle b, u_i \rangle)$$

$$M_a M_b^{-1} = V D_a D_b^{-1} V^{-1}$$

(if $r < d$, V^{-1} is "pseudo inverse")

$$(M_a^{-1} M_b)^T = W D_b^{-1} D_a W^{-1}$$

w.p. 1 $D_a(i,i) \neq 0$, $D_b(i,i) \neq 0$, $\frac{D_a(i,i)}{D_b(i,i)}$ unique

$\Longrightarrow$ v_i's are eigenvectors of $M_a M_b^{-1}$

w_i's ——————— $(M_a^{-1} M_b)^T$

and the pairing is correct.

Final step: Let $Z \in \mathbb{R}^{d^2 \times r} = \begin{pmatrix} v_1 \otimes w_1 & v_2 \otimes w_2 & \cdots & v_r \otimes w_r \end{pmatrix}$

$$= V \odot W$$

(Khatri-Rao product)

then $T_{1,(2,3)} = U Z^T$.

solution U is unique if Z is full rank.

(but this is trivial because even V is full rank) $\square$