Lect 3 and remarks about ALS.

Recap ALS also: Thee Metric chans.

\[ T(1) ; T(2) ; T(3), \]

order 3. Suppose you know that rank is R.

- Initialize also \( B^0, C^0 \) as arrays of \( \text{dim } I_2 \times R, I_3 \times R \).

full column rank.

- Initialize \( m \geq 0 \).

\[
\begin{align*}
A^{m+1} &= T_{(1)} \left( C^m \otimes_{khr} B^m \right) \left( C^m A^T + B^m B^T \right) \\
B^{m+1} &= T_{(2)} \left( A^{m+1} \otimes_{khr} C^m \right) \left[ A^{m+1} A^T + C^m C^T \right] \\
C^{m+1} &= T_{(3)} \left( B^{m+1} \otimes_{khr} A^{m+1} \right) \left( B^{m+1} B^T + A^{m+1} A^T \right)
\end{align*}
\]

This also comes with various correcs & problems

(although it is still very useful & popular).

Remarks.
Remarks:

1) Attempt is to minimize $\| T - \sum_{r=1}^{R} z_r b_r c_{r'}^{T} \|^2_E$ by min on $a^r, b^r, c^r$.

Highly non-convex and if the edge vanishes, then it might end up at some local min.

Might be a good one or not depend on initialization.

2) The rank $R$ of $T$ has to be known.

3) Because at each iteration (except first step)

There is no guarantee that enough inner products exist

$+\text{MP pseudoinverse exist always but not necessarily compatible from}$

If $A^+, C^+$ are full col rank.

$$\left( (A^{m+1} \otimes_m C^n) (A^{m+1} A^{m+1} + C C^T)^{-1} \right)$$

$$\left( (A^{m+1} \otimes_m C^n) \right)^T$$

exists.
people regularize the formulas:

$$(A^m \otimes_{khr} C^n) \left( A^{mT} A^m + C^n C^n + \lambda I \right)^{-1}$$

has $\lambda > 0$ small reg parameter

and $(\text{Identity})_{R \times R} = I$.

- Regularized algorithm makes sense certainly for $\lambda > 0$ through all iterations.

I implemented with small reg $\lambda > 0$.

4) Then, what is the quantity that we attempt to minimize in principle?

$$\| T - \sum_{r=1}^{R} b_r \otimes b_r \otimes c_r \|_F^2 + \lambda \left( \sum_{r=1}^{R} \| a_r \|_2^2 + \| b_r \|_2^2 + \| c_r \|_2^2 \right)$$

Ridge regularizer.
\[ \text{argmin}_A \left\{ \| T_{c1} - A (C \otimes_{\text{Khr}} B) \|_F^2 + 2 \| A \|_F^2 \right\} \]

\[ \text{argmin}_B \left\{ \| T_{c2} - B (A \otimes_{\text{Khr}} C) \|_F^2 + 2 \| B \|_F^2 \right\} \]

\[ \text{argmin}_C \left\{ \| T_{c3} - C (B \otimes_{\text{Khr}} A) \|_F^2 + 2 \| C \|_F^2 \right\} \]

→ Regularized version of ALS.

\[ \text{Rank } R \text{ has still to be known.} \]

\[ \text{still not guaranteed that you get true min} \]

\[ \| T - \sum \hat z_i e_i \otimes e_i \|_F^2 + \sum \| z_i \|_F^2 \]

\[ \text{still not convex} \]

→ But both believed numerically.

\[ \text{end of ALS.} \]

\[ \text{Next time: New Ranks. New Help.} \]