{step1_generator.md}

Given a d-way tensor $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ such that the data is unaligned (meaning the tensor $\mathcal{T}$ has missing entries), we consider the problem of computing a CP decomposition of rank r where some modes are infinite-dimensional and constrained to be in a Reproducing Kernel Hilbert Space (RKHS). We want to solve this using an alternating optimization approach, and our question is focused on the mode-k subproblem for an infinite-dimensional mode. For the subproblem, then CP factor matrices $A_1, \ldots, A_{k-1}, A_{k+1}, \ldots, A_d$ are fixed, and we are solving for A_k.

Our notation is as follows. Let $N = \prod_i n_i$ denote the product of all sizes. Let $n \equiv n_k$ be the size of mode k, let $M = \prod_{i \neq k} n_i$ be the product of all dimensions except k, and assume $n \ll M$. Since the data are unaligned, this means only a subset of $\mathcal{T}$'s entries are observed, and we let $q \ll N$ denote the number of observed entries. We let $T \in \mathbb{R}^{n \times M}$ denote the mode-k unfolding of the tensor $\mathcal{T}$ with all missing entries set to zero. The vec operations creates a vector from a matrix by stacking its columns, and we let $S \in \mathbb{R}^{N \times q}$ denote the selection matrix (a subset of the $N \times N$ identity matrix) such that $S^T \text{vec}(T)$ selects the q known entries of the tensor $\mathcal{T}$ from the vectorization of its mode-k unfolding. We let $Z = A_d \odot \cdots \odot A_{k+1} \odot A_{k-1} \odot \cdots \odot A_1 \in \mathbb{R}^{M \times r}$ be the Khatri-Rao product of the factor matrices corresponding to all modes except mode k. We let $B = TZ$ denote the MTTKRP of the tensor $\mathcal{T}$ and Khatri-Rao product Z. We assume $A_k = KW$ where $K \in \mathbb{R}^{n \times n}$ denotes the psd RKHS kernel matrix for mode k. The matrix W of size $n \times r$ is the unknown for which we must solve. The system to be solved is

$$\left[(Z \otimes K)^T SS^T (Z \otimes K) + \lambda(I_r \otimes K) \right] \text{vec}(W) = (I_r \otimes K) \text{vec}(B). \qquad (1)$$

Here, I_r denotes the $r \times r$ identity matrix. This is a system of size $nr \times nr$ Using a standard linear solver costs $O(n^3 r^3)$, and explicitly forming the matrix is an additional expense.

Explain how an iterative preconditioned conjugate gradient linear solver can be used to solve this problem more efficiently. Explain the method and choice of preconditioner. Explain in detail how the matrix-vector products are computed and why this works. Provide complexity analysis. We assume $n, r < q \ll N$. Avoid any computation of order N.

PCG for the RKHS mode subproblem in CP tensor completion

The mode-k normal equations take the form $\mathcal{A} \text{vec}(W) = \mathbf{b}$ of size $nr \times nr$, where $\mathcal{A} = (Z \otimes K)^T SS^T (Z \otimes K) + \lambda(I_r \otimes K)$ is symmetric positive definite (since $K \succ 0$ and $\lambda > 0$), making CG applicable. The key observation is that the matrix–vector product

$\mathcal{A} \text{vec}(W)$ can be evaluated in $O(qr + n^2 r)$ without any $O(N)$ computation. This relies on two facts: (i) the vec-Kronecker identity $(Z \otimes K) \text{vec}(W) = \text{vec}(KWZ^T)$ reduces the forward map to a dense $n \times n$ multiply $P = KW$ followed by pointwise evaluation of PZ^T at only the q observed index pairs (i_s, j_s), costing $O(n^2 r + qr)$ instead of the $O(nMr)$ needed to form the full $n \times M$ matrix; and (ii) the adjoint $(Z^T \otimes K)(Sy) = \text{vec}(KVZ)$ involves the sparse matricization $V \in \mathbb{R}^{n \times M}$ (with q nonzeros), so VZ is a scatter operation costing $O(qr)$, followed by another $O(n^2 r)$ multiply with K.

For the preconditioner, we replace the diagonal sampling operator SS^T by its expectation under uniform observation, $\frac{q}{N} I_N$, which restores pure Kronecker structure: $\mathcal{P} = \frac{q}{N}(Z^T Z \otimes K^2) + \lambda(I_r \otimes K)$. This is simultaneously diagonalized by the eigenbases of K ($n \times n$) and $Z^T Z$ ($r \times r$), reducing preconditioner solves to a change of basis and entrywise division at cost $O(n^2 r + nr^2)$ after a one-time $O(n^3 + r^3)$ eigendecomposition. Since $n, r \ll q$, each PCG iteration is dominated by the $O(qr)$ sparse operations, giving total cost $O(n^3 + r^3 + tqr)$ for t iterations—compared to $O(n^3 r^3)$ for a direct solve.