A Formally Verified PCG Framework for RKHS-Constrained Mode-k CP Updates with Missing Tensor Entries

Abstract

We study the mode-k linear subproblem arising in alternating optimization for CP decomposition with missing tensor entries and RKHS constraints. The unknown factor is parameterized as $A_k = KW$, where $K \in \mathbb{R}^{n\times n}$ is a kernel Gram matrix and $W \in \mathbb{R}^{n\times r}$ is the optimization variable. The normal equation has size $nr \times nr$ and is prohibitively expensive to solve by direct dense methods. We present an operator-based preconditioned conjugate gradient (PCG) scheme that avoids any $\mathcal{O}(N)$ work with $N = \prod_i n_i$. We derive explicit matrix–vector application formulas from observed entries only, establish per-iteration complexity bounds, and prove strict improvement over ambient-size enumeration when $q < nM$ (q: observed entries, $M = \prod_{i\neq k} n_i$). We also prove PCG admissibility under structured assumptions (symmetry and positivity of data and kernel-induced operators, plus $\lambda > 0$). The full proof chain is formalized and machine-checked in Lean 4/mathlib (no `sorry`, successful build), and we map each mathematical statement to the corresponding formal theorem.

1 Introduction

Consider a d-way tensor $\mathcal{T} \in \mathbb{R}^{n_1\times\cdots\times n_d}$ with missing entries. In alternating CP updates, fixing all factors except mode k yields a linear subproblem for the mode-k factor. In RKHS-constrained variants, one writes

$$A_k = KW,$$

where $K \in \mathbb{R}^{n\times n}$ is a positive semidefinite Gram matrix ($n = n_k$), and $W \in \mathbb{R}^{n\times r}$ is the unknown.

With the standard notation

$$N = \prod_{i=1}^{d} n_i, \qquad M = \prod_{i\neq k} n_i, \qquad q \ll N,$$

and with $Z \in \mathbb{R}^{M\times r}$ the Khatri–Rao product of the fixed factors, the mode-k normal equation is

$$\Big[(Z\otimes K)^T SS^T (Z\otimes K) + \lambda(I_r\otimes K)\Big]\mathrm{vec}(W) = (I_r\otimes K)\mathrm{vec}(B), \tag{1}$$

where $B = TZ$, T is the mode-k unfolding with missing entries filled by zero, and S selects observed entries.

A direct dense solve costs $\mathcal{O}(n^3r^3)$, and explicitly forming the coefficient matrix is itself expensive. Our goal is an iterative solver with observed-entry complexity.

Contributions.

1. We derive an operator form whose application uses only observed entries.
2. We provide explicit per-application and per-PCG-iteration complexity formulas.
3. We prove SPD-based PCG admissibility from structured assumptions.
4. We provide a machine-checked Lean 4 formalization of the proof chain.

2 Problem Setup and Observation-Driven Operator

Let observed pairs be indexed by

$$\Omega = \{(i_p, m_p)\}_{p=1}^{q} \subseteq [n] \times [M].$$

For $X \in \mathbb{R}^{n\times r}$, define

$$s_p(X) := \sum_{t=1}^{r} (KX)_{i_p t} Z_{m_p t}. \tag{2}$$

Define the data term $D(X) \in \mathbb{R}^{n\times r}$ by

$$[D(X)]_{j\ell} := \sum_{p=1}^{q} K_{j i_p}\, Z_{m_p \ell}\, s_p(X). \tag{3}$$

Then define the system operator

$$A(X) := D(X) + \lambda K X. \tag{4}$$

The right-hand side is $\mathrm{RHS} = KB$.

Proposition 1 (Operator realization of (1)). *For all $X \in \mathbb{R}^{n\times r}$,*

$$\mathrm{vec}(A(X)) = \Big[(Z \otimes K)^T S S^T (Z \otimes K) + \lambda (I_r \otimes K)\Big] \mathrm{vec}(X).$$

Hence solving (1) is equivalent to solving $A(W) = KB$.

Proof. By the mixed-product and vectorization identities,

$$\mathrm{vec}(KX) = (I_r \otimes K)\mathrm{vec}(X).$$

The selection matrix S restricts to observed entries. Expanding $(Z \otimes K)^T S S^T (Z \otimes K)\mathrm{vec}(X)$ entrywise yields exactly (2)–(3). Therefore the first term equals $\mathrm{vec}(D(X))$, and adding the regularizer gives $\mathrm{vec}(A(X))$. □

3 Matrix–Vector Product Complexity

The key implementation principle is to apply $A(\cdot)$ without forming any $nr \times nr$ matrix.

Computation pattern

For input X:

1. Compute $Y = KX$ (cost $n^2 r$ arithmetic operations).
2. For each $p \in \{1, \dots, q\}$, compute $s_p(X)$ via (2) (cost $2r$ in the adopted operation model).
3. Accumulate (3) over (j, ℓ) (cost nr per observation, i.e. qnr total).
4. Add λY.

   This yields the formal cost identity below.

Theorem 1 (Per-application cost). *In the operation model above,*

$$C_{\mathrm{mv}} = q(2r + nr) + n^2 r. \tag{5}$$

Moreover, with ambient size $N = nM$,

$$C_{\mathrm{mv}} \le nM(2r + nr) + n^2 r, \tag{6}$$

and if $q < nM$, then

$$C_{\mathrm{mv}} < nM(2r + nr) + n^2 r. \tag{7}$$

Proof. The decomposition of work above gives (5) directly. Since $q \le nM$, multiplying by $(2r + nr) > 0$ gives (6); strict inequality follows when $q < nM$. □

4 PCG and Preconditioner Theory

We work on $\mathbb{R}^{n\times r}$ with Frobenius inner product

$$\langle X, Y\rangle_F := \sum_{i=1}^{n}\sum_{j=1}^{r} X_{ij}Y_{ij}.$$

Definition 1 (Symmetry and positivity)**.** For a linear operator $\mathcal{L} : \mathbb{R}^{n\times r} \to \mathbb{R}^{n\times r}$:

$$\begin{aligned} \mathcal{L} \text{ is symmetric} &\iff \langle \mathcal{L}(X), Y\rangle_F = \langle X, \mathcal{L}(Y)\rangle_F \ \forall X, Y, \\ \mathcal{L} \text{ is positive definite} &\iff \langle \mathcal{L}(X), X\rangle_F > 0 \ \forall X \neq 0, \\ \mathcal{L} \text{ is SPD} &\iff \text{symmetric and positive definite.} \end{aligned}$$

Choose the block preconditioner

$$P_\mu := I_r \otimes (K + \mu I_n), \qquad \mu > 0, \tag{8}$$

whose matrix action is

$$M_\mu(X) = (K + \mu I_n)X = KX + \mu X.$$

We take $\mu = \lambda$.

Assumption 1 (Structured operator hypotheses)**.** *(A1) D is symmetric:*

$$\langle D(X), Y\rangle_F = \langle X, D(Y)\rangle_F.$$

(A2) D is positive semidefinite: $\langle D(X), X\rangle_F \geq 0$ for all X.

(A3) K-multiplication operator $X \mapsto KX$ is symmetric.

(A4) K-multiplication operator is positive definite.

(A5) $\lambda > 0$.

Theorem 2 (System operator is SPD)**.** *Under (A1)–(A5), the operator $A(X) = D(X) + \lambda KX$ is SPD.*

Proof. Symmetry: D is symmetric by (A1), $X \mapsto KX$ is symmetric by (A3), scalar multiplication preserves symmetry, and sums of symmetric operators are symmetric.

Positive definiteness: for $X \neq 0$,

$$\langle A(X), X\rangle_F = \langle D(X), X\rangle_F + \lambda\langle KX, X\rangle_F.$$

By (A2), the first term is nonnegative. By (A4), $\langle KX, X\rangle_F > 0$. By (A5), $\lambda > 0$. Therefore $\langle A(X), X\rangle_F > 0$. □

Theorem 3 (Default preconditioner is SPD)**.** *Under (A3)–(A5), $M_\lambda(X) = (K + \lambda I_n)X$ is SPD.*

Proof. Symmetry follows from symmetry of $X \mapsto KX$ and identity. For $X \neq 0$,

$$\langle M_\lambda(X), X\rangle_F = \langle KX, X\rangle_F + \lambda \|X\|_F^2.$$

By (A4), the first term is positive; by (A5), the second is nonnegative and strict when $X \neq 0$. Hence positivity holds. □

Corollary 1 (PCG admissibility)**.** *Under (A1)–(A5), PCG with operator A and preconditioner M_λ is admissible.*

Proof. By Theorems 2 and 3, both operators are SPD, which is the standard admissibility condition for PCG. □

5 Per-Iteration and Total Complexity

Besides one application of A, one PCG step includes a preconditioner solve/apply and vector updates. Using the verified operation model:

$$
\begin{aligned}
C_{\text{prec}} &= n^2 r, && (9)\\
C_{\text{vec}} &= 6nr, && (10)\\
C_{\text{iter}} &= C_{\text{mv}} + C_{\text{prec}} + C_{\text{vec}} \\
&= q(2r + nr) + 2n^2 r + 6nr. && (11)
\end{aligned}
$$

Therefore, for k PCG iterations,

$$
C_{\text{total}}(k) = k\, C_{\text{iter}}. \tag{12}
$$

No step requires iterating over all $N = \prod_i n_i$ tensor entries.

6 Lean Formalization Map

The formal development is in Lean 4/mathlib and compiles with no placeholders.

Mathematical statement	Lean theorem/file
Observation-driven operator and dimensions	`Question10/Defs.lean`
Matvec cost formula (5)	`costSystemMatVec_eq` in `SparseMatVec.lean`
Upper/strict bound (6),(7)	`costSystemMatVec_le_ambient_scaled`, `..._lt_...`
System SPD from structured assumptions	`spd_systemOp_of_assumptions` in `PCG.lean`
Preconditioner SPD	`spd_preconditioner_of_kernel_assumptions` in `PCG.lean`
PCG admissibility	`pcg_ready_fully_structured` in `PCG.lean`
Main bundled theorem	`main_solver_statement` in `Main.lean`
Per-iteration and total cost identities	`costPCGIter_eq`, `totalCost_eq` in `Complexity.lean`

7 Conclusion

We provided a complete operator-level analysis and formal verification pipeline for the RKHS-constrained mode-k CP subproblem with missing entries. The method avoids ambient-size computations, supports PCG with an SPD block preconditioner, and yields explicit complexity bounds per operator application and per iteration. The end-to-end argument is both mathematically explicit and machine-checked in Lean, giving a reproducible proof artifact suitable for rigorous computational mathematics workflows.