Modified definition, proof and simulation.

Chanwoo Lee

Nov 26, 2019

1 Estimation accuracy for randomized SVD

Our goal is to estimate a low-rank tensor signal tensor from a noisy tensor observation. Let $\mathcal{A} \in \mathbb{R}^{d_1 \times d_2 \times d_3}$ be the signal tensor to be of interest. We assume the signal tensor admits a low-rank Tucker decomposition,

$$\mathcal{A} = \mathcal{C} \times_1 M^{(1)} \times_2 M^{(2)} \times_3 M^{(3)}$$

where $\mathcal{C} \in \mathbb{R}^{r_1 \times r_2 \times r_3}$ and $M^{(1)}, M^{(2)}, M^{(3)}$. are orthonormal matrices in $\mathbb{R}^{d_1 \times r_1}, \mathbb{R}^{d_2 \times r_2}, \mathbb{R}^{d_3 \times r_3}$ respectively. We will suggest various methods to recover $\mathcal{A}$ from a noisy tensor

$$\mathcal{D} = \mathcal{A} + \mathcal{E}.$$

where $\mathcal{E}$ is a noise tensor with i.i.d. entries from $N(0, \sigma^2)$.

We provide some theorems to guarantee the convergence of estimators to true signal tensor in each method. Theorem 1 and Theorem 2 show that the convergence of a tensor is ensured by the convergence of orthonormal matrices. The only difference between the two theorems is distance measure used for orthonormal matrices: Theorem 1 uses Frobenius norm while Theorem 2 uses angle to measure the error. Theorem 3 suggests sufficient conditions to recover a signal tensor from a noisy observation. From these theorems, the proposed algorithms enjoy asymptotic consistency. In section 1.1, we provide the theorems to derive a generic error bound. We guarantee consistency of estimation under some conditions. The section 1.2 presents some linear algebraic tools we need in section 1.1.

1.1 Theoretical results

We develop general error bounds for suggested algorithmic methods. The first two results show that the accuracy of tensor estimation is guaranteed by the accuracy of estimations for orthonormal matrices. The last result guarantees the consistency of estimation under some conditions.

Theorem 1. *Let* $\mathcal{A} = C \times_1 M^{(1)} \times_2 M^{(2)} \cdots \times_N M^{(N)} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}$, *where* $\mathcal{C} \in \mathbb{R}^{r_1 \times r_2 \times \cdots \times r_N}$ *is a core tensor,* $\{M^{(i)} \in \mathbb{R}^{d_i \times r_i}\}$ *are the factor matrices with orthonormal columns. Let* $\hat{M}^{(i)}$ *be an estimator for* $M^{(i)}$ *such that* $\|M^{(i)} - \hat{M}^{(i)}\|_F \leq \epsilon$ *for all* $i \in [N]$. *Define* $\hat{\mathcal{A}} = \hat{\mathcal{C}} \times_1 \hat{M}^{(1)} \times_2 \hat{M}^{(2)} \cdots \times_N \hat{M}^{(N)}$ *where* $\hat{C} = \mathcal{A} \times_1 (\hat{M}^{(1)})^T \times_2 (\hat{M}^{(2)})^T \cdots \times_N (\hat{M}^{(N)})^T$. *Then the following holds,*

$$\|\hat{\mathcal{A}} - \mathcal{A}\|_F \leq \|\mathcal{A}\|_F (2 \sum_{i=1}^{N} \sqrt{r_i} + N\epsilon)\epsilon.$$

Furthermore, if we have $\|M^{(i)} - \hat{M}^{(i)}\|_\sigma \leq \epsilon$ *for all* $i \in [N]$ *instead of Frobenius norm bound, the following holds.*

$$\|\hat{\mathcal{A}} - \mathcal{A}\|_F \leq \|\mathcal{A}\|_F (2 \sum_{i=1}^{N} \|M^{(i)}\|_\sigma + N\epsilon)\epsilon.$$

Proof. First, assume the Frobenius norm error bound. Then, we have,

$$\begin{aligned}
\|M^{(i)}(M^{(i)})^T - \hat{M}^{(i)}(\hat{M}^{(i)})^T\|_F &= \|M^{(i)}(M^{(i)})^T - M^{(i)}(\hat{M}^{(i)})^T + M^{(i)}(\hat{M}^{(i)})^T - \hat{M}^{(i)}(\hat{M}^{(i)})^T\|_F \\
&= \|M^{(i)}((M^{(i)})^T - (\hat{M}^{(i)})^T) + (M^{(i)} - \hat{M}^{(i)})(\hat{M}^{(i)})^T)\|_F \\
&\leq (2\|M^{(i)}\|_F + \epsilon)\epsilon \quad \text{for all } i \in [N].
\end{aligned}$$

Let us define $P_A \coloneqq AA^T$. Then, the above inequality can be denoted as,

$$\|P_{M^{(i)}} - P_{\hat{M}^{(i)}}\|_F \leq (2\|M^{(i)}\|_F + \epsilon)\epsilon = (2\sqrt{r_i} + \epsilon)\epsilon. \quad (1)$$

The theorem's main inequality is obtained from mathematical induction.

$$
\begin{aligned}
&\|\mathcal{A}-\hat{\mathcal{A}}\|_F \\
&= \|\mathcal{A}-\hat{\mathcal{C}} \times_1 \hat{M}^{(1)} \times_2 \hat{M}^{(2)} \cdots \times_N \hat{M}^{(N)}\|_F \\
&= \|\mathcal{A}-\mathcal{A} \times_1 P_{\hat{M}^{(1)}} \times_2 P_{\hat{M}^{(2)}} \cdots \times_N P_{\hat{M}^{(N)}}\|_F \\
&= \|P_{M^{(1)}} A_{(1)} (P_{M^{(N)}} \otimes P_{M^{(N-1)}} \otimes \cdots \otimes P_{M^{(2)}})^T - P_{\hat{M}^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{\hat{M}^{(N-1)}} \otimes \cdots \otimes P_{\hat{M}^{(2)}})^T\|_F \\
&= \|P_{M^{(1)}} A_{(1)} (P_{M^{(N)}} \otimes P_{M^{(N-1)}} \otimes \cdots \otimes P_{M^{(2)}})^T - P_{M^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{M^{(N-1)}} \otimes \cdots \otimes P_{M^{(2)}})^T \\
&+ P_{M^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{M^{(N-1)}} \otimes \cdots \otimes P_{M^{(2)}})^T - P_{M^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{\hat{M}^{(N-1)}} \otimes \cdots \otimes P_{M^{(2)}})^T \\
&+ \cdots \\
&+ P_{M^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{\hat{M}^{(N-1)}} \otimes \cdots \otimes P_{\hat{M}^{(2)}})^T - P_{\hat{M}^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{\hat{M}^{(N-1)}} \otimes \cdots \otimes P_{\hat{M}^{(2)}})^T\|_F \\
&\le \|P_{M^{(1)}} A_{(1)} (P_{M^{(N)}} \otimes P_{M^{(N-1)}} \otimes \cdots \otimes P_{M^{(2)}})^T - P_{M^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{M^{(N-1)}} \otimes \cdots \otimes P_{M^{(2)}})^T\|_F \\
&+ \|P_{M^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{M^{(N-1)}} \otimes \cdots \otimes P_{M^{(2)}})^T - P_{M^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{\hat{M}^{(N-1)}} \otimes \cdots \otimes P_{M^{(2)}})^T\|_F \\
&+ \cdots \\
&+ \|P_{M^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{\hat{M}^{(N-1)}} \otimes \cdots \otimes P_{\hat{M}^{(2)}})^T - P_{\hat{M}^{(1)}} A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{\hat{M}^{(N-1)}} \otimes \cdots \otimes P_{\hat{M}^{(2)}})^T\|_F \\
&= \|(P_{M^{(1)}} - P_{\hat{M}^{(1)}}) A_{(1)} (P_{\hat{M}^{(N)}} \otimes P_{\hat{M}^{(N-1)}} \otimes \cdots \otimes P_{\hat{M}^{(2)}})^T\|_F \\
&+ \sum_{n=2}^{N} \|P_{M^{(1)}} A_{(1)} \big(P_{\hat{M}^{(N)}} \otimes \cdots \otimes P_{\hat{M}^{(n+1)}} \otimes (P_{\hat{M}^{(n)}} - P_{M^{(n)}}) \otimes P_{M^{(n-1)}} \otimes \cdots \otimes P_{M^{(2)}}\big)^T\|_F \\
&\le \|(P_{M^{(1)}} - P_{\hat{M}^{(1)}})\|_F \|A_{(1)}\|_F \prod_{i=2}^{N} \|P_{\hat{M}^{(i)}}\|_F \\
&+ \sum_{n=2}^{N} \left(\|P_{M^{(1)}}\|_F \|A_{(1)}\|_F \|P_{\hat{M}^{(n)}} - P_{M^{(n)}}\|_F \prod_{i=n+1}^{N} \|P_{\hat{M}^{(i)}}\|_F \prod_{i=2}^{n-1} \|P_{M^{(i)}}\|_F \right).
\end{aligned} \tag{2}
$$

Proposition 1 and Proposition 3 are used in the last inequality in (2). The following inequality is from Proposition 2 and the last inequality in (2).

$$
\begin{aligned}
\|\mathcal{A}-\hat{\mathcal{A}}\|_F &\le \|P_{M^{(1)}} - P_{\hat{M}^{(1)}}\|_F \|A_{(1)}\|_F + \sum_{i=2}^{N} \|P_{\hat{M}^{(i)}} - P_{M^{(i)}}\|_F \|A_{(1)}\|_F \\
&= \sum_{i=1}^{N} \|P_{\hat{M}^{(i)}} - P_{M^{(i)}}\|_F \|A_{(1)}\|_F \\
&= \sum_{i=1}^{N} \|P_{\hat{M}^{(i)}} - P_{M^{(i)}}\|_F \|\mathcal{A}\|_F
\end{aligned}
$$

Finally, we get the inequality of the theorem from (1).

Likewise, we can get the similar result under the assumption of the spectral norm bound. By the same approach for Frobenius norm case, we get,

$$\|M^{(i)}(M^{(i)})^T - \hat{M}^{(i)}(\hat{M}^{(i)})^T\|_\sigma \leq (2\|M^{(i)}\|_\sigma + \epsilon)\epsilon. \tag{3}$$

From the inequality (2), we have,

$$\begin{aligned}
\|\mathcal{A} - \hat{\mathcal{A}}\|_F &\leq \|(P_{M^{(1)}} - P_{\hat{M}^{(1)}})A_{(1)}(P_{\hat{M}^{(N)}} \otimes P_{\hat{M}^{(N-1)}} \otimes \cdots \otimes P_{\hat{M}^{(2)}})^T\|_F \\
&+ \sum_{n=2}^{N} \|P_{M^{(1)}}A_{(1)}\big(P_{\hat{M}^{(N)}} \otimes \cdots \otimes P_{\hat{M}^{(n+1)}} \otimes (P_{\hat{M}^{(n)}} - P_{M^{(n)}}) \otimes P_{M^{(n-1)}} \otimes \cdots \otimes P_{M^{(2)}}\big)^T\|_F \\
&\leq \|(P_{M^{(1)}} - P_{\hat{M}^{(1)}})\|_\sigma \|A_{(1)}\|_F \prod_{i=2}^{N} \|P_{\hat{M}^{(i)}}\|_F \\
&+ \sum_{n=2}^{N} \left(\|P_{M^{(1)}}\|_F \|A_{(1)}\|_F \|P_{\hat{M}^{(n)}} - P_{M^{(n)}}\|_\sigma \prod_{i=n+1}^{N} \|P_{\hat{M}^{(i)}}\|_F \prod_{i=2}^{n-1} \|P_{M^{(i)}}\|_F \right) \\
&\leq \|P_{M^{(1)}} - P_{\hat{M}^{(1)}}\|_\sigma \|A_{(1)}\|_F + \sum_{i=2}^{N} \|P_{\hat{M}^{(i)}} - P_{M^{(i)}}\|_\sigma \|A_{(1)}\|_F \\
&= \sum_{i=1}^{N} \|P_{\hat{M}^{(i)}} - P_{M^{(i)}}\|_\sigma \|A_{(1)}\|_F \\
&= \sum_{i=1}^{N} \|P_{\hat{M}^{(i)}} - P_{M^{(i)}}\|_\sigma \|\mathcal{A}\|_F.
\end{aligned} \tag{4}$$

Proposition 3 is used in in the second inequality. Proposition 2 and 3 are used in the last inequality. Finally, we have the desired result using (3). □

We define the principal angle to measure the distance between two subspaces.

Definition 1. *For nonzero subspaces $\mathcal{F}, \mathcal{G} \subset \mathbb{R}^n$, the principal angle between $\mathcal{F}$ and $\mathcal{G}$ is defined to be the number $0 \leq \theta \leq \pi/2$ that satisfies*

$$\sin\theta(\mathcal{F}, \mathcal{G}) = \max_{u \in \mathcal{F}^\perp, v \in \mathcal{G}, \|u\| = \|v\| = 1} v^t u.$$

Where $\mathcal{F}^\perp$ is the orthogonal complement of $\mathcal{F}$. A principal angle between two matrices is *defined as*

$$\sin\theta(A,B) = \sin\theta(\text{span}(A), \text{span}(B)),$$

where $A \in \mathbb{R}^{n\times m}, B \in \mathbb{R}^{n\times k}$ *are two matrices.*

If we use the principal angle to measure the error, the following Theorem 2 holds.

Theorem 2. *Under the same condition as in Theorem 1, let us assume that the following error bound condition holds true instead of Frobenius norm and spectral norm.*

$$\sin(\theta(M^{(i)}, \hat{M}^{(i)})) < \epsilon.$$

Then, we have

$$\|\mathcal{A} - \hat{\mathcal{A}}\|_F \leq 2\epsilon N \|\mathcal{A}\|_F.$$

Proof. If the inequality $\|P_{M^{(i)}} - P_{\hat{M}^{(i)}}\|_\sigma \leq 2\epsilon$ is true, the main inequality holds from (4) in the proof of Theorem 1.

$$\|\mathcal{A} - \hat{\mathcal{A}}\|_F \leq (\sum_{i=1}^{N} \|P_{M^{(i)}} - P_{\hat{M}^{(i)}}\|_\sigma)\|\mathcal{A}\|_F \leq 2\epsilon N\|\mathcal{A}\|_F.$$

So it suffices to show $\|P_{M^{(i)}} - P_{\hat{M}^{(i)}}\|_\sigma \leq 2\epsilon$ under the given condition.

$$\begin{aligned}
&\|P_{M^{(i)}} - P_{\hat{M}^{(i)}}\|_\sigma \\
&= \|M^{(i)}(M^{(i)})^T - \hat{M}^{(i)}(\hat{M}^{(i)})^T\|_\sigma \\
&= \|M^{(i)}(M^{(i)})^T - M^{(i)}(M^{(i)})^T\hat{M}^{(i)}(\hat{M}^{(i)})^T + M^{(i)}(M^{(i)})^T\hat{M}^{(i)}(\hat{M}^{(i)})^T - \hat{M}^{(i)}(\hat{M}^{(i)})^T\|_\sigma \\
&\leq \|M^{(i)}(M^{(i)})^T - M^{(i)}(M^{(i)})^T\hat{M}^{(i)}(\hat{M}^{(i)})^T\|_\sigma + \|M^{(i)}(M^{(i)})^T\hat{M}^{(i)}(\hat{M}^{(i)})^T - \hat{M}^{(i)}(\hat{M}^{(i)})^T\|_\sigma \\
&= \|M^{(i)}(M^{(i)})^T(I - \hat{M}^{(i)}(\hat{M}^{(i)})^T)\|_\sigma + \|(M^{(i)}(M^{(i)})^T - I)\hat{M}^{(i)}(\hat{M}^{(i)})^T\|_\sigma \\
&= \sin(\theta) + \sin(\theta) \leq 2\epsilon.
\end{aligned}$$

The last equality follows from Proposition 4. □

1.2 Some background from linear algebra

We collect linear algebraic background in this section.

Proposition 1 (Kronecker Product Norms). *Let* $\|\cdot\|_F$ *be Frobenious norm and* $\|\cdot\|_\sigma$ *be spectral norm. For any* $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times k}$

$$\|A\otimes B\|_F = \|A\|_F\|B\|_F, \quad \|A\otimes B\|_\sigma = \|A\|_\sigma\|B\|_\sigma.$$

Proof. From the definition of Frobenius norm and Kronecker product, we can get

$$\|A\otimes B\|_F = \| \begin{bmatrix} a_{11}B & \cdots & a_{mn}B \\ \vdots & \cdots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \|_F = \|A\|_F\|B\|_F.$$

For the case of spectral norm, Let $A = \sum_i \sigma_i u_i v_i^T$ and $B = \sum_j \lambda_j x_j y_j^T$ be the singular value decomposition of two matrices. Then,

$$\begin{aligned} A\otimes B &= \sum_{i,j} \sigma_i\lambda_j (u_i v_i^T)\otimes(x_j y_j^T) \\ &= \sum_{i,j} \sigma_i\lambda_j (u_i\otimes x_j)(v_i^T\otimes y_j^T) \\ &= \sum_{i,j} \sigma_i\lambda_j (u_i\otimes x_j)(v_i\otimes y_j)^T. \end{aligned}$$

From this form, we can see $\{\sigma_i\lambda_j : i,j\}$ are singular values of $A\otimes B$. Therefore,

$$\|A\otimes B\|_\sigma = \max_{i,j} \sigma_i\lambda_j = (\max_i \sigma_i)(\max_j \lambda_j) = \|A\|_\sigma\|B\|_\sigma.$$

□

Proposition 2 (Norm of Projection Matrices). *Let* $P \in \mathbb{R}^{m\times n}$ *be a projection matrix and* $A \in \mathbb{R}^{n\times k}$ *be an arbitrary matrix. Then,*

$$\|PA\|_F \le \|A\|_F, \quad \|PA\|_\sigma \le \|A\|_\sigma.$$

Proof. Notice for $x \in \mathbb{R}^n, \quad \|x\|_2^2 = \|Px\|_2^2 + \|(I-P)x\|_2^2$. Therefore, we have the following inequality

$$\|Px\|_2 \le \|x\|_2.$$

This proves the main inequalities of proposition 2. □

Proposition 3 (Frobenius Norm of Matrix Product)**.** *Let $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times k}$ be arbitrary matrices. Then*

$$\|AB\|_F \leq \|A\|_F\|B\|_F, \quad \|AB\|_F \leq \|A\|_F\|B\|_\sigma.$$

Proof. The submultiplicity of Frobenius norm is an application of Cauchy-Schwarz inequality. So we focus on proving $\|AB\|_F \leq \|A\|_F\|B\|_\sigma$ here. Let A_i denote i-th row of A. Then,

$$\|AB\|_F^2 = \sum_i \|(AB)_i\|_\sigma^2 = \sum_i \|A_iB\|_\sigma^2 \leq \sum_i \|A_i\|_F^2\|B\|_\sigma^2 = \|A\|_F^2\|B\|_\sigma^2.$$

□

The following 2 propositions offer a way to relate the principal angle and spectral norm.

Proposition 4. *Let $P_\mathcal{F}$ and $P_\mathcal{G}$ be the orthogonal projectors from $\mathbb{R}^d$ onto subspaces $\mathcal{F}$ and $\mathcal{G}$, respectively, then,*

$$\|P_\mathcal{F}(I - P_\mathcal{G})\|_\sigma = \sin(\theta).$$

Proof.

$$\begin{aligned}\|P_\mathcal{F}(I - P_\mathcal{G})\|_\sigma &= \max_{u,v\in\mathbb{R}^d,\|u\|=\|v\|=1} u^T P_\mathcal{F}(I - P_\mathcal{G})v = \max_{u,v\in\mathbb{R}^d,\|u\|\leq 1,\|v\|\leq 1} u^T P_\mathcal{F}(I - P_\mathcal{G})v \\ &= \max_{x\in P_\mathcal{F}, y\in P_\mathcal{G}^\perp,\|x\|\leq 1,\|y\|\leq 1} x^T y = \sin(\theta).\end{aligned}$$

□

Proposition 5 (Spectral Norm of Gaussin Matrix)**.** *Let $E \in R^{m\times n}$ be a random matrix with i.i.d. $N(0, 1)$ entries. Then, we have, with very high probability,*

$$\|E\|_\sigma \asymp (2 + o(1))\sqrt{\max(m, n)}.$$

2 The modified simulation result.

2.1 Simulation.

I compared 3 various randomized Tucker decomposition algorithms:

1. Method 1: Unstructured Gaussian test matrix + Single Random Projection.
2. Method 2: Khatri-Rao Gaussian test matrix + Single Random Projection.
3. Method 3: Khatri-Rao Gaussian test matrix + Multiple Random Projections.

This simulation investigates the accuracy of estimators in terms of angles and MSE for an arbitrary rank. We consider an order-3 dimension $(20, 20, 20)$ signal tensor X. We assume X has Tucker decomposition as $X = \mathcal{C} \otimes_1 B_1 \otimes_2 B_2 \otimes_3 B_3$, where $B_i \in \mathbb{R}^{20\times 3}$ for all i. and $\mathcal{C} \in \mathbb{R}^{3\times 3\times 3}$ a core tensor. All entries of $\mathcal{C}$ are i.i.d. drawn from $\mathrm{N}(0, 1)$. B_1, B_2, B_3 are randomly drawn from orthonormal matrices. We vary the noise level $\sigma \in \{0.001, 0.002, \cdots 0.024, 0.025\}$. We use target rank 3 and estimate the signal matrices according to each algorithms. We compare the principal angles between the true signal matrices and estimators. Figure 1 shows that Method 3 outperforms the other methods in all respects.

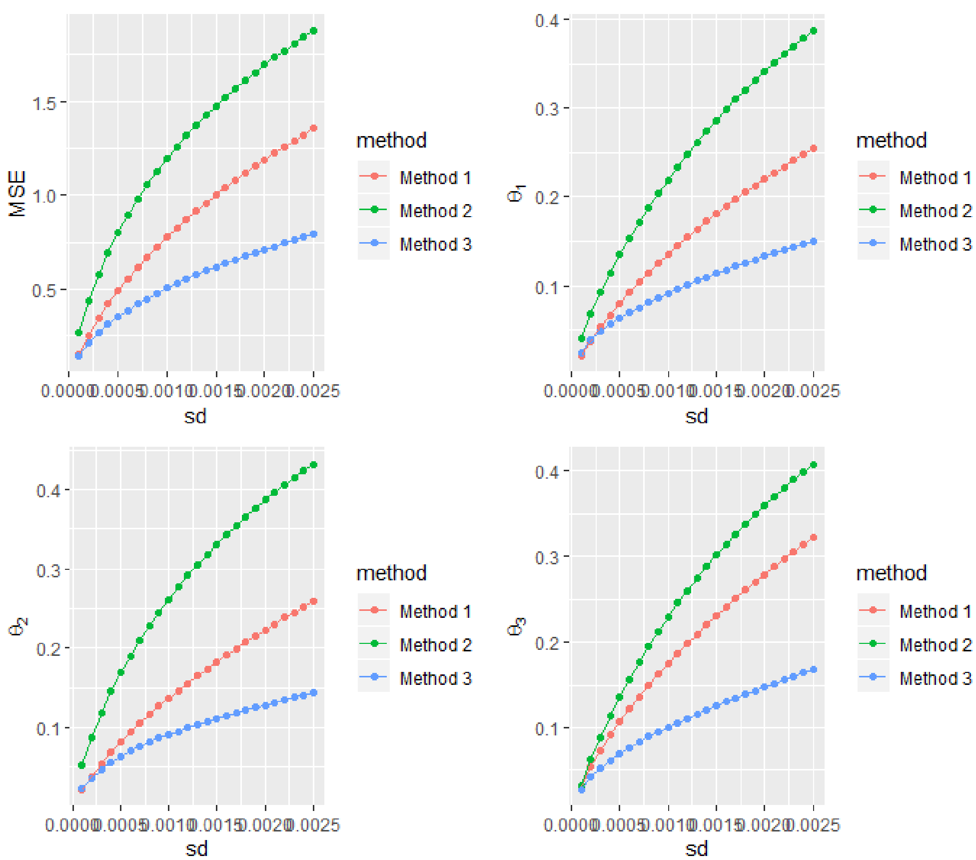

Figure 1: The y axis represents MSE and the priciple angles between the true signal and estimator. The x axis represents the noise size.

3 Simulation code.

```
B_1 = randortho(20,type=c("orthonormal"))[,1:3]
B_2 = randortho(20,type=c("orthonormal"))[,1:3]
B_3 = randortho(20,type=c("orthonormal"))[,1:3]
C = as.tensor(array(rnorm(3^3),dim = c(3,3,3)))
X = ttm(ttm(ttm(C,B_1,1),B_2,2),B_3,3)
sd = 0.0001*1:25
result = data.frame(matrix(0,nrow = 75, ncol =6))
names(result) <- c("sd","angle1","angle2","angle3","method","MSE")



for (i in 1:25) {
  s=sd[i]
  result[i,1] = s
  result[i+25,1] = s
  result[i+50,1] = s
  for (j in 1:200) {
    set.seed(j)
    e = as.tensor(array(rnorm(8000,mean =0,sd = s),dim = c(20,20,20)))
    D = X+e
    est1 = tensor_svd(D,3,3,3,0)
    est2 = tensor_svd3(D,3,3,3,0)
    est3 = tensor_svd4(D,3,3,3,0)
    result[i,2] <- result[i,2]+angle(est1$U[[1]],B_1)
    result[i,3] <- result[i,3]+angle(est1$U[[2]],B_2)
    result[i,4] <- result[i,4]+angle(est1$U[[3]],B_3)
    result[i,6] <- result[i,6]+tensor_resid(X,est1)
    result[i+25,2] <- result[i+25,2]+angle(est2$U[[1]],B_1)
    result[i+25,3] <- result[i+25,3]+angle(est2$U[[2]],B_2)
    result[i+25,4] <- result[i+25,4]+angle(est2$U[[3]],B_3)
    result[i+25,6] <- result[i+25,6]+tensor_resid(X,est2)
    result[i+50,2] <- result[i+50,2]+angle(est3$U[[1]],B_1)
    result[i+50,3] <- result[i+50,3]+angle(est3$U[[2]],B_2)
    result[i+50,4] <- result[i+50,4]+angle(est3$U[[3]],B_3)
    result[i+50,6] <- result[i+50,6]+tensor_resid(X,est3)
  }
  result[i,5] = "Method 1"
```

```
  result[i+25,5] = "Method 2"
  result[i+50,5]= 'Method 3'
}
result[,2:4] <- result[,2:4]/200
result[,6] <- result[,6]/200
r3 <- result
```