Supplementary Notes for
Tucker-Structured Phase Retrieval

Soo Min Kwon

October 1, 2021

The objective of this supplementary set of notes is to show how we can use Conjugate Gradient Least Squares (CGLS) to update each factor matrix for the algorithm Tucker-Structured Phase Retrieval.

Preliminaries

Vectorization: Define $\mathbf{X} \in \mathbb{C}^{n_1 \times n_2}$. The vec$(\cdot)$ operator creates a column (or row) vector from any multi-dimensional array by stacking the columns (or rows) of the array. For example, if we let

$$\mathbf{x} = \text{vec}(\mathbf{X}), \tag{1}$$

then $\mathbf{x}$ has dimensions n (i.e. $\mathbf{x} \in \mathbb{C}^n$), where $n = n_1 n_2$. There are many important properties of the vec$(\cdot)$ operator [1]. The property that we will use is that given matrices $\mathbf{A} \in \mathbb{C}^{q \times n_1}$ and $\mathbf{B} \in \mathbb{C}^{n_2 \times r}$,

$$\text{vec}(\mathbf{AXB}) = (\mathbf{B}^\top \otimes \mathbf{A})\text{vec}(\mathbf{X}), \tag{2}$$

where $\otimes$ is the Kronecker product. Consequently, we can see that

$$\begin{aligned} \text{vec}(\mathbf{XB}) &= \text{vec}(\mathbf{IXB}) \qquad (3)\\ &= (\mathbf{B}^\top \otimes \mathbf{I})\text{vec}(\mathbf{X}) \qquad (4) \end{aligned}$$

and

$$\begin{aligned} \text{vec}(\mathbf{AX}) &= \text{vec}(\mathbf{AXI}) \qquad (5)\\ &= (\mathbf{I}^\top \otimes \mathbf{A})\text{vec}(\mathbf{X}), \qquad (6) \end{aligned}$$

where $\mathbf{I}$ is the identity matrix.

Least Squares: Recall that a least squares problem can posed as an optimization problem of the form

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}^\top \mathbf{x}\|^2, \tag{7}$$

where $\mathbf{A} \in \mathbb{C}^{n\times m}$, $\mathbf{y} \in \mathbb{C}^m$, and $\mathbf{x} \in \mathbb{C}^n$. The optimal $\mathbf{x}$, denoted as $\mathbf{x}^*$, has the closed-form solution

$$\mathbf{x}^* = (\mathbf{A}^\top \mathbf{A})^{-1}\mathbf{A}^\top \mathbf{y}. \tag{8}$$

This is equivalent to saying that the optimization formulation is solving the linear system

$$\mathbf{A}^\top \mathbf{A}\mathbf{x} = \mathbf{A}^\top \mathbf{y}. \tag{9}$$

If we can rearrange our optimization formula to look in the form of equation (9), then we can use conjugate gradient least squares (CGLS) to solve for the optimal $\mathbf{x}^*$.

Tucker-Structured Phase Retrieval

Notation: All norms (e.g. $\|\cdot\|$) are ℓ_2-norms unless otherwise stated. Scalars are denoted with lowercase letters (e.g. x), vectors are denoted as bold lowercase letters (e.g. $\mathbf{x}$), matrices are denoted as bold uppercase letters (e.g. $\mathbf{X}$), and tensors are denoted as underlined bold letters (e.g. $\underline{\mathbf{X}}$). Matricization of order d will be denoted as $\mathcal{M}_d(\cdot)$. Since we are working in the complex domain, all matrix transposes can be regarded as Hermitian (or conjugate) transposes.

Problem Formulation: Recall that in the setup of Tucker-Structured Phase Retrieval, we assume that our tensor $\underline{\mathbf{X}} \in \mathbb{C}^{n_1\times n_2\times q}$ admits a Tucker decomposition of the form

$$\underline{\mathbf{X}} = \underline{\mathbf{G}} \times_1 \mathbf{D} \times_2 \mathbf{E} \times_3 \mathbf{F}, \tag{10}$$

where $\underline{\mathbf{G}} \in \mathbb{C}^{r_1\times r_2\times r_3}$, $\mathbf{D} \in \mathbb{C}^{n_1\times r_1}$, $\mathbf{E} \in \mathbb{C}^{n_2\times r_2}$, and $\mathbf{F} \in \mathbb{C}^{q\times r_3}$. We want to solve for these factor matrices (and core tensor) given sampling matrices $\mathbf{A}_k$ and observations

$$y_{i,k} = |\langle \mathbf{a}_{i,k}, \mathrm{vec}(\mathbf{X}_k)\rangle|^2, \tag{11}$$

for $k = 1,\dots,q$ and $i = 1,\dots,m$, where $\mathbf{X}_k \in \mathbb{C}^{n_1\times n_2}$ are the frontal slices of $\underline{\mathbf{X}} \in \mathbb{C}^{n_1\times n_2\times q}$. The general optimization formula for alternating minimization in phase retrieval is

$$\sum_k \|\mathbf{C}_k\mathbf{y}_k - \mathbf{A}_k\mathrm{vec}(\mathbf{X}_k)\|_2^2. \tag{12}$$

To solve for the Tucker factors with this formulation, we need to rewrite $\mathbf{x}_k = \mathrm{vec}(\mathbf{X}_k)$ in terms of the Tucker factors and core tensor.

Updating D: When updating $\mathbf{D}$, we can write our objective function as

$$\sum_k \left\|\mathbf{C}_k\sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}(\mathbf{X}_k)\right\|^2 = \sum_k \left\|\mathbf{C}_k\sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}(\mathbf{D}\cdot\mathcal{M}_1(\underline{\mathbf{G}})(\mathbf{f}_k\otimes\mathbf{E})^\top)\right\|^2. \tag{13}$$

Let $\mathbf{S}_k = \mathcal{M}_1(\underline{\mathbf{G}})(\mathbf{f}_k\otimes\mathbf{E})^\top$. Then, our objective function becomes

$$\begin{aligned}\sum_k \left\|\mathbf{C}_k\sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}(\mathbf{D}\mathbf{S}_k)\right\|^2 &= \sum_k \left\|\mathbf{C}_k\sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}(\mathbf{I}\mathbf{D}\mathbf{S}_k)\right\|^2 &(14)\\ &= \sum_k \left\|\mathbf{C}_k\sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top (\mathbf{S}_k^\top\otimes\mathbf{I})\mathrm{vec}(\mathbf{D}))\right\|^2, &(15)\end{aligned}$$

where the second equality comes from the property previously stated. Lastly, if we let

$$\mathbf{T}_k = \mathbf{A}_k^\top (\mathbf{S}_k^\top \otimes \mathbf{I}), \tag{16}$$

updating $\mathbf{D}$ amounts to solving the system

$$\mathbf{T}_k^\top \mathbf{T}_k \mathrm{vec}(\mathbf{D}) = \mathbf{T}_k^\top \mathbf{C}_k \sqrt{\mathbf{y}_k}. \tag{17}$$

Upon solving for $\mathrm{vec}(\mathbf{D})$, we can reshape the vector back into $\mathbf{D} \in \mathbb{C}^{n_1 \times r_1}$.

Updating E: When updating $\mathbf{E}$, we can write our objective function as

$$\sum_k \left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}(\mathbf{X}_k) \right\|^2 = \sum_k \left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}((\mathbf{E} \cdot \mathcal{M}_2(\underline{\mathbf{G}})(\mathbf{f}_k \otimes \mathbf{D})^\top)^\top) \right\|^2. \tag{18}$$

Let $\mathbf{U}_k = \mathcal{M}_2(\underline{\mathbf{G}})(\mathbf{f}_k \otimes \mathbf{D})^\top$. Then,

$$\sum_k \left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}((\mathbf{I}\mathbf{E}\mathbf{U}_k)^\top) \right\|^2. \tag{19}$$

Using the same property, the optimization problem becomes

$$\sum_k \left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}((\mathbf{I}\mathbf{E}\mathbf{U}_k)^\top) \right\|^2 = \sum_k \left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}(\mathbf{U}_k^\top \mathbf{E}^\top \mathbf{I}^\top) \right\|^2 \tag{20}$$

$$= \sum_k \left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top (\mathbf{I} \otimes \mathbf{U}_k^\top) \mathrm{vec}(\mathbf{E}^\top) \right\|^2. \tag{21}$$

With $\mathbf{V}_k \coloneqq \mathbf{A}_k^\top (\mathbf{I} \otimes \mathbf{U}^\top)$, we can now solve the linear system

$$\mathbf{V}_k^\top \mathbf{V}_k \mathrm{vec}(\mathbf{E}^\top) = \mathbf{V}_k^\top \mathbf{C}_k \sqrt{\mathbf{y}_k}. \tag{22}$$

Updating F: The update step for factor matrix $\mathbf{F}$ is slightly different in the sense that we have to solve for each column of $\mathbf{F}$, $\mathbf{f}_k$, separately. In addition, we do not use least squares to solve for $\mathbf{f}_k$, and instead use Reshaped Wirtinger Flow to solve a noisy r-dimensional phase retrieval problem. However, we can still take the methods shown previously to create the sampling matrices needed for Reshaped Wirtinger Flow. The objective function can be written as

$$\left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}(\mathbf{f}_k \cdot \mathcal{M}_3(\underline{\mathbf{G}})(\mathbf{E} \otimes \mathbf{D})^\top) \right\|^2. \tag{23}$$

Let $\mathbf{H} = \mathcal{M}_3(\underline{\mathbf{G}})(\mathbf{E} \otimes \mathbf{D})^\top$. The formulation is now

$$\left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top \mathrm{vec}(\mathbf{I}\mathbf{f}_k\mathbf{H}) \right\|^2 = \left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top (\mathbf{H}^\top \otimes \mathbf{I}) \mathrm{vec}(\mathbf{f}_k) \right\|^2 \tag{24}$$

$$= \left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{J}_k^\top \mathrm{vec}(\mathbf{f}_k) \right\|^2, \tag{25}$$

where $\mathbf{J}_k = \mathbf{A}_k^\top (\mathbf{H}^\top \otimes \mathbf{I})$. We can solve for $\mathbf{f}_k$ using Reshaped Wirtinger Flow with the assumption that $\mathbf{f}_k$ were generated by

$$y_k = |\mathbf{J}_k^\top \mathbf{f}_k|. \tag{26}$$

Updating $\mathcal{G}$: Lastly, when updating $\underline{\mathbf{G}}$, we can write our least squares objective function as

$$\sum_k \left\| \mathbf{C}_k \sqrt{\mathbf{y}_k} - \mathbf{A}_k^\top (\mathbf{f}_k \otimes \mathbf{E} \otimes \mathbf{D}) \mathrm{vec}(\underline{\mathbf{G}})) \right\|^2. \tag{27}$$

With $\mathbf{M}_k \coloneqq \mathbf{A}_k^\top (\mathbf{D} \otimes \mathbf{E} \otimes \mathbf{f}_k)$,

$$\mathbf{M}_k^\top \mathbf{M}_k \mathrm{vec}(\underline{\mathbf{G}}) = \mathbf{M}_k^\top \mathbf{C}_k \sqrt{\mathbf{y}_k}. \tag{28}$$

We can solve for $\mathrm{vec}(\underline{\mathbf{G}})$ and reshape it back into its tensor form.

References

[1] K. B. Petersen and M. S. Pedersen, "The matrix cookbook," Nov 2012, version 20121115. [Online]. Available: http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html