Convergence Analysis of “Blind Image Deblurring Using Row-Column Sparse Representations”

Mohammad Tofighi, Yuelong Li, Vishal Monga

In the following analysis, we assume that $r$ is sufficiently large so that all the factorizations over $X = ZH^T$ are well defined; a conservative bound (depending only on $K, N, L$) for $r$ may be obtained using the methods discussed in [1], etc. We also assume $\mathcal{A}$ is an operator satisfying the following property: $\|A(X)\|_F^2 \geq \delta \|X\|_F^2$ for some $\delta > 0, \forall X \in \mathbb{R}^{K \times N}$ and rank $(X) \leq 2r$. This property is a weaker form of the restricted isometry property (RIP) discussed in [2], and has been employed in many previous works such as [3]. We further assume that L-BFGS succeeds in finding a local minima $(Z^{k+1}, H^{k+1})$ to $L_{\sigma, \mathcal{A}}(Z, H, w_k; \alpha_k), \ \forall k$.

We will prove the convergence of Algorithm 1 that solves (5); proof corresponding to the algorithm solving (6) can be derived analogously. For ease of analysis (and to compensate for the scaling between (5) and (7) in the paper), we let $f^*$ be twice of the optimal cost of (5) in paper and let $X^* = Z^*H^{-T}$ be its minimizer. We further define $f(Z, H, w) := \|Z\|_F^2 + \|H\|_F^2 + \sum_{k=1}^K \left(w_i + \lambda^2 \frac{\|Z^T H_i\|_2^2}{w_i}\right)$. We also define $w^*$ via $w^*_i = \lambda \|Z^*H_i\|_2/2$.

**Proposition 1.** Every local minima $(\hat{Z}, \hat{H})$ of $L_{\sigma, \mathcal{A}}(Z, H, w; \alpha)$ globally minimizes the Lagrangian $L_{\sigma, \mathcal{A}}(Z, H, w; \hat{\alpha})$, where $\hat{\alpha} = \alpha - \sigma(A(ZH^T) - \hat{y})$.

**Proof.** Following the same reasoning as in Proposition 2.3 in [1], $(\hat{Z}, \hat{H})$ minimizes the following problem through the mapping $X = ZH^T, V = ZZ^T, \text{and } W = HH^T$:

$$
\min_{X, W, V} \text{tr}(V) + \text{tr}(W) + \lambda^2 \sum_{i=1}^K \frac{\|e_i^T X\|_2^2}{w_i} + 2\langle \alpha, \hat{y} - A(X) \rangle + \sigma \|A(X) - \hat{y}\|^2
$$

subject to $\begin{bmatrix} V & X \\ X^T & W \end{bmatrix} \succeq 0$, and in turn minimizes the following dual problem by the mapping $X = ZH^T$:

$$
\min_{X} 2\|X\|_* + \lambda^2 \sum_{i=1}^K \frac{\|e_i^T ZH^T\|_2^2}{w_i} + 2\langle \alpha, \hat{y} - A(X) \rangle + \sigma \|A(X) - \hat{y}\|^2.
$$

Therefore, $0 \in \partial \|ZH^T\|_* + \lambda^2 \text{diag}(w)^{-1} ZH^T - A^*(\hat{\alpha})$ where $\partial$ denotes the subdifferential and $A^*$ is the adjoint operator to $A$. This implies:

$$
\|ZH^T\|_* + \frac{\lambda^2}{2} \sum_{i=1}^K \frac{\|e_i^T ZH^T\|_2^2}{w_i} - \langle A^*(\hat{\alpha}), ZH^T \rangle \leq \|X\|_* + \frac{\lambda^2}{2} \sum_{i=1}^K \frac{\|e_i^T ZH^T\|_2^2}{w_i} - \langle A^*(\hat{\alpha}), X \rangle \ \forall X \in \mathbb{R}^{K \times N}.
$$

In particular, letting $X = ZH^T$ gives

$$
\|ZH^T\|_* + \frac{\lambda^2}{2} \sum_{i=1}^K \frac{\|e_i^T ZH^T\|_2^2}{w_i} - \langle A^*(\hat{\alpha}), ZH^T \rangle \leq \|ZH^T\|_* + \frac{\lambda^2}{2} \sum_{i=1}^K \frac{\|e_i^T ZH^T\|_2^2}{w_i} - \langle A^*(\hat{\alpha}), ZH^T \rangle \ \forall H \in \mathbb{R}^{N \times r}, Z \in \mathbb{R}^{K \times r}.
$$

On the other hand, since $\nabla_Z L_{\sigma, \mathcal{A}}(Z, H, w; \alpha) = 0, \ \nabla_H L_{\sigma, \mathcal{A}}(Z, H, w; \alpha) = 0$:

$$
\begin{align*}
\hat{Z} - A^*(\hat{\alpha}) \hat{H} + \lambda^2 \text{diag}(w)^{-1} ZH^T \hat{H} &= 0, \\
\hat{H}^T - ZH^T A^*(\hat{\alpha}) + \lambda^2 ZH^T \text{diag}(w)^{-1} ZH^T &= 0.
\end{align*}
$$

Left multiplying (5) by $Z^T$, right multiplying (6) by $H$ and subtracting the two gives $Z^T Z = H^T H$; thus $\hat{Z}, \hat{H}$ admits the following singular value decomposition:

$$
\hat{Z} = U_1 S V^T, \hat{H} = U_2 S V^T, U_1 \in \mathbb{R}^{K \times r}, U_2 \in \mathbb{R}^{N \times r}, S \in \mathbb{R}^{r \times r}, V \in \mathbb{R}^{r \times r}.
$$
and therefore
\[ \|ZH^T\|_* = \|U_1S^2U^T\|_* = \sum_{i=1}^r S_i = \frac{1}{2} \left( \|Z\|_F^2 + \|H\|_F^2 \right) \]
where \( S = \begin{bmatrix} S_1 & & \\ & \ddots & \\ & & S_r \end{bmatrix} \).

Following the same procedures on \( \nabla_Z \mathcal{L}_{0,\lambda}(Z^*, H^*, w^*; \alpha^*) = 0 \) and \( \nabla_H \mathcal{L}_{0,\lambda}(Z^*, H^*, w^*; \alpha^*) = 0 \), we can obtain
\[ \|Z^*H^T\|_* = \frac{1}{2} \left( \|Z^*\|_F^2 + \|H^*\|_F^2 \right). \]  

On the other hand, using the Arithmetic-Geometric Mean Inequality [4]:
\[ \sum_i s_i(ZH^T) \leq \frac{1}{2} \sum_i \left[ s_i(Z^T Z)^2 + s_i(H^T H)^2 \right] = \frac{1}{2} \left( \|Z\|_F^2 + \|H\|_F^2 \right), \] where \( s_i(X) \) is the \( i \)-th singular value of \( X \). Plugging (8) and (10) into (4) gives \( \mathcal{L}_{0,\lambda}(Z, H, w; \hat{\alpha}) \leq \mathcal{L}_{0,\lambda}(Z, H, w; \hat{\alpha}) \).

**Proposition 2.** \( \|w^k - w^*\|_2^2 \leq 2 \left( \lambda^2\|Z^kH^kT - Z^*H^*T\|_F^2 + \frac{\lambda^2}{(k+2)^2} \right) \).

**Proof.** Using the mean-value theorem
\[ |w^k - w^*| = \left\| \lambda e_i^T Z^k H^kT \right\|_2 + \frac{\varepsilon_0}{(k+2)^2} \leq \lambda \left\| e_i^T Z^k H^kT \right\|_2 + \frac{\varepsilon_0}{(k+2)^2} \]
\[ \leq \lambda \left\| (Z^k H^kT - Z^* H^*T) e_i \right\|_2 + \frac{\varepsilon_0}{(k+2)^2} \] (Cauchy-Schwarz),

where \( Z_\theta = \theta Z^* + (1 - \theta) Z^k, H_\theta = \theta H^* + (1 - \theta) H^k, \theta \in (0, 1) \). Therefore
\[ \|w^k - w^*\|_2^2 = \sum_{i=1}^K |w_i^k - w_i^*|^2 \leq \sum_{i=1}^K 2 \left( \lambda^2 \left\| (Z^k H^kT - Z^* H^*T) e_i \right\|_2 + \frac{\varepsilon_0^2}{(k+2)^4} \right) = 2 \left( \lambda^2 \left\| Z^k H^kT - Z^* H^*T \right\|_F^2 + \frac{\varepsilon_0^2}{(k+2)^4} \right). \]

**Proposition 3.** \( f^* \leq \mathcal{L}_{0,\lambda}(Z, H, w; \alpha^*), \forall Z \in \mathbb{R}^{K \times r}, H \in \mathbb{R}^{N \times r}, \) and \( w \in \mathbb{R}^K \).

**Proof.** Since \( X^* = Z^* H^{*T} \) minimizes the convex program (5) in paper, it also minimizes its Lagrangian, and we thus have
\[ f^* = \|Z^* H^{*T}\|_* + \lambda \|Z^* H^{*T}\|_{2,1} \leq \|X\|_* + \lambda \|X\|_{2,1} + \langle \alpha^*, \tilde{y} - A(X) \rangle, \forall X \in \mathbb{R}^{K \times N}. \] (12)

By taking \( X = ZH^T \), the RHS becomes
\[ \|ZH^T\|_* + \lambda \|ZH^T\|_{2,1} + \langle \alpha^*, \tilde{y} - A(ZH^T) \rangle \leq \frac{1}{2} \left( \|Z\|_F^2 + \|H\|_F^2 \right) + \sum_{i=1}^K \left( w_i + \lambda^2 \frac{\|e_i^T ZH^T\|_2^2}{w_i} \right) + \langle \alpha^*, \tilde{y} - A(ZH^T) \rangle \]
\[ = \frac{1}{2} \mathcal{L}_{0,\lambda}(Z, H, w; \alpha^*), \] (13)
where we used the inequality \( \|ZH^T\|_* \leq \frac{1}{2} \left( \|Z\|_F^2 + \|H\|_F^2 \right) \) (as in the proof of the Proposition 1) and \( w_i + \lambda^2 \frac{\|e_i^T ZH^T\|_2^2}{w_i} \geq 2\sqrt{\lambda^2 w_i \|e_i^T ZH^T\|_2^2} = 2\lambda \|e_i^T ZH^T\|_2 \).
Theorem 1. The sequence \((Z^k, H^k)\) generated by Algorithm 1 converges to the optimal solution of (5) in the sense of \(\lim_{k \to \infty} Z^k H^{kT} = Z^* H^{*T} = X^*\).

Proof. Let \(\gamma^k = \tilde{y} - A(Z^k H^{kT})\). By Proposition 1, \(\mathcal{L}_{0,A}(Z^{k+1}, H^{k+1}, w^k; \alpha^{k+1}) \leq \mathcal{L}_{0,A}(Z^*, H^*, w^k; \alpha^{k+1})\), i.e.
\[
f(Z^{k+1}, H^{k+1}, w^k) + \langle \alpha^{k+1}, \gamma^{k+1} \rangle \leq f(Z^*, H^*, w^k).
\]
(14)

By Proposition 3, \(f^* \leq \mathcal{L}_{0,A}(Z^{k+1}, H^{k+1}, w^k; \alpha^*)\), i.e.
\[
f^* \leq f(Z^{k+1}, H^{k+1}, w^k) + \langle \alpha^{k+1}, \gamma^{k+1} \rangle.
\]
(15)

Adding (14) and (15) gives
\[
\langle \alpha^{k+1} - \alpha^*, \gamma^{k+1} \rangle \leq f(Z^*, H^*, w^k) - f^*.
\]
(16)

Recall \(\alpha^k = \alpha^{k+1} - \sigma_k \gamma^{k+1}\); thus
\[
\|\alpha^k - \alpha^*\|_2^2 - \|\alpha^{k+1} - \alpha^*\|_2^2 = \sigma_k^2 \|\gamma^{k+1}\|_2^2 - 2 \sigma_k \langle \alpha^{k+1} - \alpha^*, \gamma^{k+1} \rangle \
\geq \sigma_k^2 \|\gamma^{k+1}\|_2^2 - 2 \sigma_k \|f(Z^*, H^*, w^k) - f^*\| (\text{using (16)})
\]
(17)

On the other hand, since \(Z^k H^{kT} - Z^* H^{*T}\) has rank at most \(2r\), we may invoke the RIP property of \(A\) to get
\[
\|\gamma^{k}\|_2^2 = \|A(Z^k H^{kT} - Z^* H^{*T})\|_2^2 \geq \delta \|Z^k H^{kT} - Z^* H^{*T}\|_F^2,
\]
(18)

and by letting \(\varepsilon_k = \frac{\varepsilon_0}{(k+1)^2}\), we have
\[
f(Z^*, H^*, w^k) - f^* = \sum_{i=1}^{k} \left( w_i^k + \frac{w_i^k w_i^k - 2 w_i^k}{w_i^k} \right) (\text{using (9)})
\]
\[
= \sum_{i=1}^{k} \left( 1 - \frac{(w_i^k)^2}{w_i^k} \right) \leq \frac{\|w^k - w^*\|_2^2}{2 \varepsilon_k} \quad (\text{since } w_i^k \geq \varepsilon_k)
\leq \frac{\lambda^2}{\varepsilon_k^2} \|Z^* H^{kT} - Z^* H^{*T}\|_F^2 + K \varepsilon_k \quad (\text{using Proposition 2})
\leq \frac{\lambda^2 (k+1)^2}{\varepsilon_0 \delta} \|\gamma^{k}\|_2^2 + \frac{K \varepsilon_0}{(k+1)^2}.
\]
(19)

Plugging (19) into (17), we get
\[
\|\alpha^k - \alpha^*\|_2^2 - \|\alpha^{k+1} - \alpha^*\|_2^2 \geq \sigma_k^2 \|\gamma^{k+1}\|_2^2 - \frac{2 \sigma_k \lambda^2 (k+1)^2}{\varepsilon_0 \delta} \|\gamma^{k}\|_2^2 - \frac{2 \sigma_k K \varepsilon_0}{(k+1)^2}.
\]
(20)

And by left division over \(\sigma_k\), we have
\[
\frac{\|\alpha^k - \alpha^*\|_2^2}{\sigma_k} - \frac{\|\alpha^{k+1} - \alpha^*\|_2^2}{\sigma_k} \geq \frac{\|\alpha^k - \alpha^*\|_2^2}{\sigma_k} - \frac{\|\alpha^{k+1} - \alpha^*\|_2^2}{\sigma_k} \geq \sigma_k \|\gamma^{k+1}\|_2^2 - \frac{2 \lambda^2 (k+1)^2}{\varepsilon_0 \delta} \|\gamma^{k}\|_2^2 - \frac{2 K \varepsilon_0}{(k+1)^2}.
\]
(21)

Since \(\sigma_k = \sigma_0 \rho^k, \sigma_k > \frac{2 \lambda^2 (k+1)^2}{\varepsilon_0 \delta} + 1\) for \(k \geq \bar{k}\). Summing over \(k\) to \(\infty\) gives
\[
\frac{\|\alpha^k - \alpha^*\|_2^2}{\sigma_k} \geq \sum_{k=k}^{\infty} \left[ \sigma_k - \frac{2 \lambda^2 (k+1)^2}{\varepsilon_0 \delta} \right] \|\gamma^{k+1}\|_2^2 - \sigma_k \frac{2 \lambda^2 (k+1)^2}{\varepsilon_0 \delta} \|\gamma^{k}\|_2^2 - \sum_{k=k}^{\infty} \frac{2 K \varepsilon_0}{(k+1)^2}
\]
\[
\geq \sum_{k=k}^{\infty} \|\gamma^{k+1}\|_2^2 - \sigma_k \frac{2 \lambda^2 (k+1)^2}{\varepsilon_0 \delta} \|\gamma^{k}\|_2^2 - \sum_{k=k}^{\infty} \frac{2 K \varepsilon_0}{(k+1)^2}.
\]
(22)

Since the left hand side is bounded and \(\sum_{k=k}^{\infty} \frac{1}{(k+1)^2}\) converges, \(\sum_{k=k}^{\infty} \sigma_k \|\gamma^{k+1}\|_2^2\) converges and thus \(\gamma^{k+1} \to 0\). Moreover, as \(\|Z^k H^{kT} - Z^* H^{*T}\|_F^2 \leq \frac{1}{\delta} \|A(Z^k H^{kT} - Z^* H^{*T})\|_F^2\), we have \(\lim_{k \to \infty} Z^k H^{kT} = Z^* H^{*T}\).

Finally, we note that in general it is impossible to prove \(\lim_{k} Z^k = Z^*\) and \(\lim_{k} H^k = H^*\) separately due to the fact that \((Z^*, H^*)\) cannot be uniquely identified. To understand this point, let \(Q \in \mathbb{R}^{r \times r}\) be any orthogonal matrix, then \((Z^* Q, H^* Q)\) also certifies as a minimizer to (7) in the paper. However, the dominating singular vector remains the same except for possible sign changes, and under the context of blind deblurring such ambiguity is inconsequential.
REFERENCES