Estimation Error for Intercept Case

Jiaxin Hu

June 29, 2021

1 Hard Constraint

Consider the optimization problem

$$\begin{aligned}
\min_{U,\Theta_r} \quad & \mathcal{L}(U,\Theta_r) = \sum_{k=1}^{K} \langle S_k, \Omega_k \rangle - \log\det(\Omega_k) \\
s.t. \quad & \Omega_k = \Theta_0 + \sum_{r=1}^{R} u_{kr}\Theta_r, \\
& \|\Theta_0\|_0 \leq s_0, \quad \|\Theta_r\|_0 \leq s_r \\
& \|u_{\cdot r}\|_F = 1, \quad \sum_{k=1}^{K} u_{kr} = 0, \quad \text{for all } r \in [R].
\end{aligned}$$

Notations.

1. Let U^*, Θ_r^*, I_r^* denote the true parameters and membership.
2. Let $I_r = \{k \in [K] : u_{kr} \neq 0\}$ collects the categories that belong to group r with given membership U, and $I_{ar} = \{k \in [K] : u_{kr}, u_{ka}^* \neq 0\}$ collects the categories that belong to group r and true group a with given membership U and the true membership U^*.
3. Let $\Sigma_k = (\Theta_0^* + u_{kr}^*\Theta_r^*)^{-1}$ be the true precision matrix for $k \in I_r^*$.
4. Let $0 < \min_{k\in[K]} \varphi_{\min}(\Sigma_k) \leq \max_{k\in[K]} \varphi_{\max}(\Sigma_k) < \tau^{-1}$.
5. Let $\Delta_0 = \Theta_0 - \Theta_0^*$, $\Delta_{ar} = \Theta_r - \Theta_a^*$, and $\Delta_{k,ar} = \Delta_0 + u_{kr}\Theta_r - u_{ka}^*\Theta_a^*$.

Lemma 1. *There exists a local minimizer for the optimization problem 1 satisfies the following inequalities simultaneously with high probability.*

$$\|\Delta_0\|_F \leq M_0\sqrt{\frac{s_0 \log p}{nK}}, \quad \|\Delta_{ar}\|_F \leq M_{ar}\sqrt{\frac{(s_r + s_a)\log p}{n|I_{ar}|}}, \quad |u_{kr} - u_{ka}^*| \leq M_k\sqrt{\frac{s_r \log p}{n}},$$

for $k \in I_{ar}, a, r \in [R]$ and some large positive constants M_0, M_{ar}, M_k.

Remark 1. The above lemma approximately agrees with the heuristic that

$$\left\|\hat{\theta}-\theta^*\right\|_F^2 = \frac{\text{degree of freedom}}{\text{sample size}}.$$

Note that there are nK samples include the intercept matrix, $|I_{rr}n|$ samples contribute to the estimation of Θ_r, and only n samples contributes to the estimation of u_{kr}. Thus, the inequality of u_{kr} may be further sharpened.

Proof. Consider the estimate (U, Θ_r) and the true parameters (U^*, Θ_r^*). Define the function

$$G(U, \Theta_r) = \mathcal{L}(U, \Theta_r) - \mathcal{L}(U^*, \Theta_r^*).$$

Note that $G(U^*, \Theta_r^*) = 0$. Therefore, our goal is to find a set $\mathcal{A}$ such that when $(U, \Theta_r) \in \partial\mathcal{A}$ we have $G(U, \Theta_r) > 0$. Thus, there exists a local minimizer inside the set $\mathcal{A}$. For simplicity, we does not consider the group only with intercept, and we assume $I_{ar} > 0$ for all $a, r \in [R]$. In next step, we may consider the group only with intercept and the case with $I_{ar} = 0$.

Rewrite the function G, we have

$$\begin{aligned}
G(U, \Theta_r) &= \sum_{r=1}^{R}\sum_{a=1}^{R}\left[\sum_{k\in I_{ar}}\langle S_k, \Delta_{k,ar}\rangle - \log\det(\Theta_0 + u_{kr}\Theta_r) + \log\det(\Theta_0^* + u_{ka}^*\Theta_a^*)\right] \\
&\geq I_1 + I_2,
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}}\langle S_k - \Sigma_k, \Delta_{k,ar}\rangle \\
I_2 &= \frac{1}{4\tau^2}\sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}}\|\Delta_{k,ar}\|_F^2.
\end{aligned}$$

For the first term, we have

$$I_1 = \langle\sum_{k\in[K]}(S_k - \Sigma_k), \Delta_0\rangle + \sum_{r=1}^{R}\sum_{a=1}^{R}\langle\sum_{k\in I_{ar}}u_{ka}^*(S_k - \Sigma_k), \Delta_{ar}\rangle + \sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}}(u_{kr} - u_{ka}^*)\langle S_k - \Sigma_k, \Theta_r\rangle.$$

By Lemma 3, with high probability, we have

$$\begin{aligned}
\left\|\sum_{k\in[K]}(S_k - \Sigma_k)\right\|_{\max} &\leq C_0\sqrt{\frac{\log pK}{n}}, \\
\left\|\sum_{k\in I_{ar}}u_{ka}^*(S_k - \Sigma_k)\right\|_{\max} &\leq C_{ar}\sqrt{\frac{\log p|I_{ar}|}{n}}, \\
\|(S_k - \Sigma_k)\|_{\max} &\leq C_k\sqrt{\frac{\log p}{n}},
\end{aligned}$$

for positive constants $C_0, C_{ar}, C_k, a, r \in [R], k \in [K]$. By the inequality $|\langle A, B\rangle| \leq \|A\|_{\max} \|B\|_1$ and the fact that $\|\Delta\|_1 \leq \sqrt{\|\Delta\|_0}\,\|\Delta\|_F$, we obtain the lower bound for I_1,

$$
\begin{aligned}
I_1 \geq & -C_0\sqrt{\frac{2s_0 \log pK}{n}}\,\|\Delta_0\|_F - \sum_{r=1}^{R}\sum_{a=1}^{R} C_{ar}\sqrt{\frac{(s_r+s_a)\log p|I_{ar}|}{n}}\,\|\Delta_{ar}\|_F \qquad (1)\\
& - \sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}} |u_{kr} - u_{ka}^*| C_k \sqrt{\frac{s_r \log p}{n}}\,\|\Theta_r\|_F
\end{aligned}
$$

For the second term, we have

$$
\|\Delta_{k,ar}\|_F^2 = \|\Delta_0\|_F^2 + \|u_{ka}^*\Delta_{ar} + (u_{kr} - u_{ka}^*)\Theta_r\|_F^2 + 2\langle \Delta_0, u_{kr}\Theta_r - u_{ka}^*\Theta_a^*\rangle.
$$

Note that

$$
\sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}} \langle \Delta_0, u_{kr}\Theta_r - u_{ka}^*\Theta_a^*\rangle = \sum_{r=1}^{R}\sum_{k\in I_r} u_{kr}\langle \Delta_0, \Theta_r\rangle - \sum_{a=1}^{R}\sum_{k\in I_a^*} u_{ka}^*\langle \Delta_0, \Theta_a^*\rangle = 0.
$$

Then, we have

$$
\begin{aligned}
I_2 &= \frac{1}{4\tau^2}\sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}} \|\Delta_0\|_F^2 + \|u_{ka}^*\Delta_{ar} + (u_{kr} - u_{ka}^*)\Theta_r\|_F^2 \qquad (2)\\
&= \frac{1}{4\tau^2}\left\{K\|\Delta_0\|_F^2 + \sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}} \left[(u_{ka}^*)^2\|\Delta_{ar}\|_F^2 + (u_{kr} - u_{ka}^*)^2\|\Theta_r\|_F^2 + 2\langle u_{ka}^*\Delta_{ar}, (u_{kr} - u_{ka}^*)\Theta_r\rangle\right]\right\},
\end{aligned}
$$

where the last term satisfies

$$
2\langle u_{ka}^*\Delta_{ar}, (u_{kr} - u_{ka}^*)\Theta_r\rangle \geq -2|u_{ka}^*||(u_{kr} - u_{ka}^*)|\,\|\Delta_{ar}\|_F\,\|\Theta_r\|_F
$$

Now consider the set

$$
\begin{aligned}
\mathcal{A} = \Bigg\{(U, \Theta_r) : \|\Delta_0\|_F \leq M_0\sqrt{\frac{s_0\log p}{nK}}, \|\Delta_{ar}\|_F \leq M_{ar}\sqrt{\frac{(s_r+s_a)\log p}{n|I_{ar}|}},\\
|u_{kr} - u_{ka}^*| \leq M_k\sqrt{\frac{s_r\log p}{n}}, k \in I_{ar}, a, r \in [R]\Bigg\},
\end{aligned}
$$

for some large constants M_0, M_{ar}, M_k. For $(U, \Theta_r) \in \partial\mathcal{A}$, we have

$$
\begin{aligned}
G(U, \Theta_r) = & \frac{M_0 s_0 \log p}{n}\left[\frac{M_0}{4\tau^2} - C_0\sqrt{2}\right] + \sum_{r=1}^{R}\sum_{a=1}^{R}\frac{M_{ar}(s_r+s_a)\log p}{n}\left[\frac{\sum_{k\in I_{ar}}(u_{ka}^*)^2 M_{ar}}{|I_{ar}|} - C_{ar}\right]\\
& + \sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}}\frac{M_k\sqrt{s_r}\log p}{n}\|\Theta_r\|_F\left[M_k\sqrt{s_r}\,\|\Theta_r\|_F - \frac{2M_{ar}|u_{ka}^*|\sqrt{s_r+s_a}}{\sqrt{|I_{ar}|}} - C_k\right].
\end{aligned}
$$

Choosing proper M_0, M_{ar}, M_k, we have $G(U, \Theta_r) > 0$, which implies these is a local minimizer lies inside $\mathcal{A}$. □

2 Soft Constraint

Consider the optimization problem

$$\begin{aligned}
\min_{U,\Theta_r} \quad & \mathcal{Q}(U,\Theta) = \sum_{k=1}^{K} \langle S_k, \Omega_k \rangle - \log\det(\Omega_k) + \lambda \left[K \left\|\Theta_0\right\|_1 + \sum_{r=1}^{R} |I_r| \left\|\Theta_r\right\|_1 \right] \\
s.t. \quad & \Omega_k = \Theta_0 + \sum_{r=1}^{R} u_{kr}\Theta_r, \\
& \left\|u_{\cdot r}\right\|_F = 1, \quad \sum_{k=1}^{K} u_{kr} = 0, \quad \text{for all } r \in [R].
\end{aligned}$$

Notations.

1. Let $\mathcal{L}(U,\Theta_r) = \sum_{k=1}^{K}\langle S_k, \Omega_k\rangle - \log\det(\Omega_k)$ denote the log-likelihood and $\mathcal{R}(U,\Theta_r) = K\left\|\Theta_0\right\|_1 + \sum_{r=1}^{R} |I_r| \left\|\Theta_r\right\|_1$ denote the penalty term.
2. Let U^*, Θ_r^*, I_r^* denote the true parameters and membership.
3. Let $I_r = \{k \in [K] : u_{kr} \neq 0\}$ collects the categories that belong to group r with given membership U, and $I_{ar} = \{k \in [K] : u_{kr}, u_{ka}^* \neq 0\}$ collects the categories that belong to group r and true group a with given membership U and the true membership U^*.
4. Let $\Sigma_k = (\Theta_0^* + u_{kr}^*\Theta_r^*)^{-1}$ be the true precision matrix for $k \in I_r^*$.
5. Let $0 < \min_{k\in[K]} \varphi_{\min}(\Sigma_k) \leq \max_{k\in[K]} \varphi_{\max}(\Sigma_k) < \tau^{-1}$.
6. Let $\Delta_0 = \Theta_0 - \Theta_0^*$, $\Delta_{ar} = \Theta_r - \Theta_a^*$, and $\Delta_{k,ar} = \Delta_0 + u_{kr}\Theta_r - u_{ka}^*\Theta_a^*$.
7. Let $s_r = \left\|\Theta_r^*\right\|_0, r = 0,1,...,R$ denote the sparsity of the true precision matrices. Let $T_r = \{(i,j) : \Theta_{r,ij}^* \neq 0\}, r = 0,1,...,R$ denote collection of nonzero entries in the true precision matrices.

Lemma 2. *Suppose*

$$\Lambda_1 \max\left\{\sqrt{\frac{\log p}{nK}}, \max_{a,r\in[R]} \sqrt{\frac{\log p}{n|I_{ar}|}}\right\} \leq \lambda \leq \Lambda_2 \min\left\{\sqrt{\frac{\log p}{nK}}, \min_{a,r\in[R]} \sqrt{\frac{\log p}{n|I_{ar}|}}\right\},$$

for some positive constants Λ_1, Λ_2. There exists a local minimizer of the soft constrained problem satisfies the following inequalities simultaneously with high probability.

$$\left\|\Delta_0\right\|_F \leq M_0\sqrt{\frac{s_0 \log p}{nK}}, \quad \left\|\Delta_{ar}\right\|_F \leq M_{ar}\sqrt{\frac{s_a \log p}{n|I_{ar}|}}, \quad |u_{kr} - u_{ka}^*| \leq M_k\sqrt{\frac{p^2 \log p}{n}},$$

for $k \in I_{ar}, a, r \in [R]$ and some large positive constants M_0, M_{ar}, M_k.

Remark 2. The Lemma 2 implies the same bounds for the hard constrained case up to some constant factors.

Proof. Consider the estimate (U, Θ_r) and the true parameters (U^*, Θ_r^*). Define the function

$$G(U, \Theta_r) = \mathcal{L}(U, \Theta_r) - \mathcal{L}(U^*, \Theta_r^*) + \lambda \left[\mathcal{R}(U, \Theta_r) - \mathcal{R}(U^*, \Theta_r^*)\right].$$

Note that $G(U^*, \Theta_r^*) = 0$. Similar with hard constraint case, our goal is to find a set $\mathcal{A}$ such that when $(U, \Theta_r) \in \partial\mathcal{A}$ we have $G(U, \Theta_r) > 0$. Thus, there exists a local minimizer inside the set $\mathcal{A}$. For simplicity, we does not consider the group only with intercept, and we assume $I_{ar} > 0$ for all $a, r \in [R]$. In next step, we may consider the group only with intercept and the case with $I_{ar} = 0$.

Rewrite the function G, we have

$$G(U, \Theta_r) \geq I_1 + I_2 + I_3,$$

where

$$\begin{aligned}
I_1 &= \sum_{r=1}^{R} \sum_{a=1}^{R} \sum_{k \in I_{ar}} \langle S_k - \Sigma_k, \Delta_{k,ar} \rangle \\
I_2 &= \frac{1}{4\tau^2} \sum_{r=1}^{R} \sum_{a=1}^{R} \sum_{k \in I_{ar}} \left\|\Delta_{k,ar}\right\|_F^2, \\
I_3 &= \lambda \left[K \left(\|\Theta_0\|_1 - \|\Theta_0^*\|_1\right) + \sum_{r=1}^{R} \sum_{a=1}^{R} |I_{ar}| \left(\|\Theta_r\|_1 - \|\Theta_a^*\|_1\right) \right].
\end{aligned}$$

For the first term, by similar procedures to obtain (1) in hard constraint case, we have

$$I_1 \geq -C_0 \sqrt{\frac{\log pK}{n}} \|\Delta_0\|_1 - \sum_{r=1}^{R} \sum_{a=1}^{R} C_{ar} \sqrt{\frac{\log p |I_{ar}|}{n}} \|\Delta_{ar}\|_1 - \sum_{r=1}^{R} \sum_{a=1}^{R} \sum_{k \in I_{ar}} |u_{kr} - u_{ka}^*| C_k \sqrt{\frac{p^2 \log p}{n}} \|\Theta_r\|_F.$$

By the inequality that $\|\Delta\|_1 = \|\Delta_T\|_1 + \|\Delta_{T^c}\|_1$, we have

$$\begin{aligned}
I_1 \geq & -C_0 \sqrt{\frac{\log pK}{n}} \left[\left\|\Delta_{0,T_0^c}\right\|_1 + \|\Delta_{0,T_0}\|_1 \right] - \sum_{r=1}^{R} \sum_{a=1}^{R} C_{ar} \sqrt{\frac{\log p |I_{ar}|}{n}} \left[\left\|\Delta_{ar,T_a^c}\right\|_1 + \|\Delta_{ar,T_a}\|_1 \right] \\
& - \sum_{r=1}^{R} \sum_{a=1}^{R} \sum_{k \in I_{ar}} |u_{kr} - u_{ka}^*| C_k \sqrt{\frac{p^2 \log p}{n}} \|\Theta_r\|_F. \qquad (3)
\end{aligned}$$

For the second term, by similar procedures to obtain (2), we have

$$I_2 \geq \frac{1}{4\tau^2} \left\{ K \|\Delta_0\|_F^2 + \sum_{r=1}^{R} \sum_{a=1}^{R} \sum_{k \in I_{ar}} \left[(u_{ka}^*)^2 \|\Delta_{ar}\|_F^2 + (u_{kr} - u_{ka}^*)^2 \|\Theta_r\|_F^2 - 2|u_{ka}^*||(u_{kr} - u_{ka}^*)| \|\Delta_{ar}\|_F \|\Theta_r\|_F \right] \right\}.$$

For the third term, by the Lemma 3 in the supplement of (Negahban et al., 2012), we have

$$I_3 \geq \lambda \left[K \left(\left\|\Delta_{0,T_0^c}\right\|_1 - \|\Delta_{0,T_0}\|_1 \right) + \sum_{r=1}^{R} \sum_{a=1}^{R} |I_{ar}| \left(\left\|\Delta_{ar,T_a^c}\right\|_1 - \|\Delta_{ar,T_a}\|_1 \right) \right]. \qquad (4)$$

Combining the inequality (3) and (4), we have

$$
\begin{aligned}
I_1 + I_3 \geq & \left(\lambda K - C_0\sqrt{\frac{\log pK}{n}}\right)\left\|\Delta_{0,T_0^c}\right\|_1 - \left(\lambda K + C_0\sqrt{\frac{\log pK}{n}}\right)\left\|\Delta_{0,T_0}\right\|_1 \\
& + \sum_{r=1}^{R}\sum_{a=1}^{R}\left(\lambda|I_{ar}| - C_{ar}\sqrt{\frac{\log p|I_{ar}|}{n}}\right)\left\|\Delta_{ar,T_a^c}\right\|_1 - \left(\lambda|I_{ar}| + C_{ar}\sqrt{\frac{\log p|I_{ar}|}{n}}\right)\left\|\Delta_{ar,T_a}\right\|_1 \\
& - \sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}}|u_{kr} - u_{ka}^*|C_k\sqrt{\frac{p^2\log p}{n}}\left\|\Theta_r\right\|_F.
\end{aligned}
$$

By the assumption that $\lambda \geq \Lambda_1 \max\left\{\sqrt{\frac{\log p}{nK}}, \max_{a,r\in[R]}\sqrt{\frac{\log p}{n|I_{ar}|}}\right\}$, we have

$$
\left(\lambda K - C_0\sqrt{\frac{\log pK}{n}}\right) > 0, \quad \left(\lambda|I_{ar}| - C_{ar}\sqrt{\frac{\log p|I_{ar}|}{n}}\right) > 0,
$$

for Λ_1 large enough. Then, with the fact that $\|\Delta\|_1 \leq \sqrt{\|\Delta\|_0}\,\|\Delta\|_F$ and the result of I_2, we have

$$
\begin{aligned}
G(U,\Theta_r) \geq & I_2 - \left(\lambda K\sqrt{s_0} + C_0\sqrt{\frac{s_0\log pK}{n}}\right)\left\|\Delta_{0,T_0}\right\|_F - \sum_{r=1}^{R}\sum_{a=1}^{R}\left(\lambda|I_{ar}|\sqrt{s_a} + C_{ar}\sqrt{\frac{s_a\log p|I_{ar}|}{n}}\right)\left\|\Delta_{ar,T_a}\right\|_F \\
& - \sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}}|u_{kr} - u_{ka}^*|C_k\sqrt{\frac{p^2\log p}{n}}\left\|\Theta_r\right\|_F.
\end{aligned}
$$

Now consider the set

$$
\begin{aligned}
\mathcal{A} = \Bigg\{(U,\Theta_r) : \|\Delta_0\|_F \leq M_0\sqrt{\frac{s_0\log p}{nK}}, \|\Delta_{ar}\|_F \leq M_{ar}\sqrt{\frac{s_a\log p}{n|I_{ar}|}}, \\
|u_{kr} - u_{ka}^*| \leq M_k\sqrt{\frac{p^2\log p}{n}}, k \in I_{ar}, a, r \in [R]\Bigg\}.
\end{aligned}
$$

For $(U,\Theta_r) \in \partial\mathcal{A}$, we have

$$
\begin{aligned}
G(U,\Theta_r) \geq & \frac{M_0 s_0\log p}{n}\left[\frac{M_0}{4\tau^2} - C_0 - \lambda\sqrt{\frac{nK}{\log p}}\right] + \sum_{r=1}^{R}\sum_{a=1}^{R}\frac{M_{ar}s_a\log p}{n}\left[\frac{M_{ar}\sum_{k\in I_{ar}}(u_{ka}^*)^2}{4\tau^2|I_{ar}|} - C_{ar} - \lambda\sqrt{\frac{n|I_{ar}|}{\log p}}\right] \\
& + \sum_{r=1}^{R}\sum_{a=1}^{R}\sum_{k\in I_{ar}}\frac{M_k p\log p}{n}\left\|\Theta_r\right\|_F\left[\frac{M_k p}{4\tau^2}\left\|\Theta_r\right\|_F - C_k p - 2|u_{ka}^*|M_{ar}\sqrt{\frac{s_a}{|I_{ar}|}}\right].
\end{aligned}
$$

By the assumption that $\lambda \leq \Lambda_2 \min\left\{\sqrt{\frac{\log p}{nK}}, \min_{a,r\in[R]} \sqrt{\frac{\log p}{n|I_{ar}|}}\right\}$, we have

$$\frac{M_0}{4\tau^2} - C_0 - \lambda\sqrt{\frac{nK}{\log p}} \geq \frac{M_0}{4\tau^2} - C_0 - \Lambda_2 > 0$$

$$\frac{M_{ar}\sum_{k\in I_{ar}}(u^*_{ka})^2}{4\tau^2|I_{ar}|} - C_{ar} - \lambda\sqrt{\frac{n|I_{ar}|}{\log p}} \geq \frac{M_{ar}\sum_{k\in I_{ar}}(u^*_{ka})^2}{4\tau^2|I_{ar}|} - C_{ar} - \Lambda_2 > 0,$$

$$\frac{M_k p}{4\tau^2}\|\Theta_r\|_F - C_k p - 2|u^*_{ka}|M_{ar}\sqrt{\frac{s_a}{|I_{ar}}} > 0$$

for small Λ_2 and large M_0, M_{ar}, M_k. Thus $G(U, \Theta_r)$ for $(U, \Theta_r) \in \partial\mathcal{A}$, and there exists a local minimizer inside $\mathcal{A}$.

□

Lemma 3. *Let* $Z_i \sim \mathcal{N}_p(\mathbf{0}, \Sigma_i)$ *i.i.d. with* $\Sigma_i = [\![\Sigma_{i,jk}]\!]$ *for* $i \in [n]$ *and* $\max_{i\in[n]} \lambda_{\max}(\Sigma_i) \leq \epsilon_0 < \infty$. *Then, we have*

$$P\left(|\frac{1}{n}\sum_{i=1}^{n}(Z_{i,j}Z_{i,k} - \Sigma_{i,jk})| \geq t\right) \leq c_1 \exp\left(-c_2 n t^2\right), \quad \textit{for} \quad t \leq |b|,$$

where c_1, c_2, b *depend on* ϵ_0.

Proof. The result follows by the equation (2.20) in (**?**). □

References

Negahban, S. N., Ravikumar, P., Wainwright, M. J., Yu, B., et al. (2012). A unified framework for high-dimensional analysis of m-estimators with decomposable regularizers. Statistical science, 27(4):538–557.