arXiv:2301.00360v2 [stat.ME] 1 Aug 2023

一种高效的迭代最小二乘算法

基于随机投影的大维矩阵因子模型

Yong He*、赵然*和Wen-Xin Zhou†

2023年8月2日

矩阵因子模型因其能够同时实现矩阵结构观测的双向降维而日益受到关注。本文针对矩阵因子模型提出了一种简单的迭代最小二乘算法，与文献中基于主成分分析（PCA）的方法形成对照。具体而言，我们首先提出利用两个确定性权重矩阵对观测值进行投影，以估计潜在因子矩阵；这两个矩阵的选择旨在消除特异性成分的影响。我们证明，即使高估行/列因子数，对因子的推断仍然渐近有效。随后，我们在一定的可识别性条件下，通过最小化平方损失函数来估计行/列载荷矩阵。所得载荷矩阵估计量被视为新的权重/投影矩阵，因此可以迭代执行上述更新步骤，直至收敛。理论方面，在给定因子矩阵真实维度的情况下，我们推导了载荷矩阵和公共成分的估计量在任意 s步迭代中的收敛速率。此外，我们提出了一种特征值比率法，用于一致地估计因子数对。我们通过全面的数值模拟考察所提方法的有限样本表现，并使用两个涉及金融投资组合和跨国宏观经济指数的真实数据集，展示该算法的实际效用。

关键词：潜在低秩；最小二乘；矩阵因子模型；随机投影。

1 引言

因子建模是高维时间序列分析中一种极为流行的降维方法，已成功应用于大规模时间序列面板，用于预测宏观经济变量（Stock 和 Watson，2002a），并构建反映整体经济

*山东大学金融研究院，中国。电子邮箱： heyong@sdu.edu.cn，Zhaoran@mail.sdu.edu.cn
†加利福尼亚大学圣迭戈分校数学科学系，美国。电子邮箱： wez243@ucsd.edu

活动的低维指标（Stock 和 Watson，2002b）。过去二十年来，关于高维因子模型的研究蓬勃发展；例如，Bai（2003年）、Onatski（2009）、Ahn and Horenstein (2013)、Fan 等人（2013年）、Trapani（2018）、Aït-Sahalia 和 Xiu（2017）、Kong（2017）、Barigozzi 等人（2018）、Yu 等人（2019）、Barigozzi 和 Cho（2020年）、Chen 等人（2021）、He 等人（2022a）以及 Fan 和 Liao（2022）。

在经济学和金融学中，观测值通常具有良好的结构，呈数组或矩阵形式，例如记录多个国家多种宏观经济变量的表格时间序列，或在线平台上大量客户对众多项目的评分序列。过去几年，文献日益关注矩阵时间序列的因子分析。 Wang 等（2019）(2019首次提出了以下矩阵时间序列因子模型:

$$\mathbf{X}_t = \mathbf{R}\mathbf{F}_t\mathbf{C}^\top + \mathbf{E}_t,\ t = 1, \ldots, T, \tag{1.1}$$

其中，R 是行因子载荷矩阵，用于提取 X 在行方向上的变化；C 是列因子载荷矩阵，反映 X 各列之间的差异；F 是公共因子矩阵，E 是特质成分。Wang 等（2019）通过对自交叉协方差矩阵进行特征值分析，提出了因子载荷矩阵的估计量；Chen 和 Fan（2021）提出了一种 α-主成分分析方法，通过对数据均值矩阵以及列（行）协方差矩阵的加权平均进行特征值分析；Yu 等（2022）提出了投影估计（PE）方法，进一步提高了因子载荷矩阵的估计效率。He等（2021a）证明了最小化平方损失与 Yu 等（2022）的 PE 方法等价，并进一步提出以 Huber 损失代替平方损失的稳健方法。对投影数据的加权样本协方差矩阵进行特征值分析，即可简单地得到因子载荷矩阵的合成估计量。He 等（2022b）提出通过对广义行／列矩阵 Kendall's tau 进行特征值分析，恢复矩阵椭圆因子模型的载荷空间；该方法将多元 Kendall tau 推广到了随机矩阵情形。然而，据我们所知，现有文献中所有关于矩阵因子模型的理论研究，都关键依赖于因子数对 k_1 和 k_2 的一致估计这一假设，而这通常要求因子相对较强、数据的序列相关较弱，或观测值数量较大。在实际应用中，由于信噪比较弱或存在非平稳性，这些条件可能无法满足，导致行／列协方差矩阵最大的若干特征值与其余特征值之间的间隔变小；另见 Fan 和 Liao（2022）关于向量因子模型的讨论。正如 Moon 和 Weidner （ 2015 ）、 Westerlund 和 Urbain （ 2015 ）以及 Barigozzi 和 Cho （ 2020 ）在经典向量因子模型中所讨论的，高估因子数量可能是一种可行的补救方法。高估矩阵因子模型中因子数对的影响仍不明确。

本文提出了一种用于矩阵因子模型的简单迭代最小二乘算法，以区别于

文献中基于主成分分析（PCA）的方法。具体而言，在第一步中，我们提出使用两个确定性权重矩阵对矩阵观测值进行投影，以估计潜在因子矩阵；这两个权重矩阵经过选择，可消除特异性成分的影响。这一思路与 Fan 和 Liao (2022) 的方法类似，而且该估计量不依赖特征向量。在第二步中，我们在可识别性条件下，通过最小化平方损失函数来更新行/列载荷矩阵。随后，将载荷矩阵的估计量作为新的权重矩阵，重复上述两个步骤，直至达到收敛准则。当前工作的贡献体现在以下几个方面。首先，据我们所知，这是首项无需对大型矩阵进行任何特征值分解的矩阵因子分析研究。所提出的迭代最小二乘算法非常简单且计算效率高，其计算复杂度为 $O(Tp_1p_2)$，而基于特征值分解的方法通常具有 $O\left(Tp_1^2+Tp_2^2\right)$ 的复杂度。其次，我们证明，即使行因子数和列因子数均被高估，对因子得分的推断仍然渐近有效，这在矩阵因子模型文献中尚属首次。第三，在因子数设定正确的情况下，我们证明，在某个强因子可识别性条件下，对于任意 s，载荷矩阵和公共成分的估计量在第 $s \geq 1$ 次迭代时的收敛速率。与一步估计（Yu 等人，2022）相比，已证明多步估计量对初始估计量的敏感性较低。此外，迭代最小二乘算法还能在每一步降低特异性误差成分的大小，从而提高信噪比，并具有与 Yu 等人(2022) 提出的投影估计方法相同的优势。最后，我们提出一种用于确定因子数量的特征值比率法，并证明该方法具有一致性。

本文其余部分安排如下。第 2 节介绍通过双向多样化投影进行因子估计的方法，并推导出即使因子数被高估，向量化因子空间仍具有一致性。第 3 节提出载荷空间的迭代最小二乘估计量，并推导出估计量在任意第 s 步迭代时的收敛速率。我们还介绍模型选择准则的一致性。第 4 节开展全面的数值研究，以展示所提方法相较于最先进方法的优势。第 5 节分析一个金融数据集和一个跨国宏观经济指数数据集，以展示所提方法的实证应用价值。我们在第 6 节讨论未来可能的研究方向并总结全文。主要定理的证明及其他补充细节汇集于补充材料中。

本节最后介绍一些将在全文中使用的记号。对于任意向量 $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_p)^\top \in \mathbb{R}^p$，令 $\|\boldsymbol{\mu}\|_2=(\sum_{i=1}^p \mu_i^2)^{1/2}$, $\|\boldsymbol{\mu}\|_\infty=\max_i |\mu_i|$。对于实数 a，记 $[a]$为小于或等于 a的最大整数。 $\otimes$ 表示克罗内克积。对于矩阵 A，令 A_{ij}(或 $A_{i,\ j}$) 表示 A 的第 (i,j) 个元素，$A^\top$ 表示 A 的转置，tr(A) 表示 A 的迹，rank(A) 表示 A 的秩，

A^+ 表示 A 的 Moore-Penrose广义逆，diag(A) 表示由 A 的对角元素组成的向量。令 Vec(A) 为将 A 的各列依次堆叠得到的向量。记 $P_{\mathbf{A}}$ 为投影矩阵 $P_{\mathbf{A}}$ =A(A$^\top$A)$^{-1}$A$^\top$ ，span(A) 表示由 A 的各列张成的空间。记 λ_j(A) 为非负定矩阵 A 的第 j 大特征值，并令 $\|A\|_2$ 表示矩阵 A 的谱范数，$\|A\|_F$ 表示 A 的 Frobenius范数。对于两列随机变量， $X_n \lesssim Y_n$ 表示 $X_n = O_p(Y_n)$， $X_n \gtrsim Y_n$ 表示 $Y_n = O_p(X_n)$。记号 X_n 和 Y_n，$X_n \asymp Y_n$ 分别表示 $X_n = O_p(Y_n)$ 和 $Y_n = O_p(X_n)$。不同公式行中的常数 c, C_1, C_2 可以各不相同。

2 通过双向多样化投影估计因子

本节介绍一种简单的双向多样化投影方法，用于估计因子得分矩阵。我们还研究了在有限样本以及因子数过度估计情况下估计量的理论性质。

2.1 因子的估计

本节提出一种简单的双向多样化投影方法，用于估计因子得分矩阵，该方法不涉及特征值分解。设 $W_i = (\boldsymbol{w}_{i,\circ 1}, \boldsymbol{w}_{i,\circ 2}, \ldots, \boldsymbol{w}_{i,\cdot m_i})$ 为给定的外生（或确定性） $p_i \times m_i$ 矩阵 $i = 1, 2$，其中 $\boldsymbol{w}_{i,\cdot j}$（矩阵 W_i 的第 j 列）是一个“多样化权重”向量，也就是说，其强度应大致均匀地分布在大多数分量上。我们分别称 W_1 和 W_2 为“左投影矩阵”和“右投影矩阵”。我们提出，只需通过以下方式估计 F_t：

$$\widehat{\mathbf{F}}_t = \frac{1}{p_1 p_2} \mathbf{W}_1^\top \mathbf{X}_t \mathbf{W}_2. \tag{2.1}$$

根据矩阵因子模型（1.1），我们有

$$\widehat{\mathbf{F}}_t = \frac{1}{p_1 p_2} \mathbf{W}_1^\top \mathbf{R} \mathbf{F}_t \mathbf{C}^\top \mathbf{W}_2 + \frac{1}{p_1 p_2} \mathbf{W}_1^\top \mathbf{E}_t \mathbf{W}_2 := \mathbf{H}_1 \mathbf{F}_t \mathbf{H}_2^\top + \boldsymbol{\mathcal{E}}_t, \tag{2.2}$$

其中 $H_1 = W_1^\top R/p_1$、$H_2 = W_2^\top C/p_2$，且 $E_t = W_1^\top E_t W_2/(p_1 p_2)$。因此，$\widehat{\ }F_t$ 对 F_t 的估计仅相差两个仿射变换，其中 E_t 为估计误差。假设 W_1,W_2 应当具有多样性，这保证了当 $\min(p_1, p_2) \to \infty$ 时，E_t 会被多样化效应消除（依概率收敛于零）。我们将新的因子矩阵估计量称为“双重多样化因子”，它通过双向多样化投影将 X_t 的维度从 $p_1 \times p_2$ 降至$m_1 \times m_2$。由于展开式 (2.2) 简洁明了，其理论分析的数学处理比大多数基准估计量简单得多。直观而言，只要 $m_i \geq k_i, i = 1, 2$，F_t 就能在因子增强模型中得到有效推断，这与 Fan 和 Liao（2022）的思路一致

对于向量因子模型，我们将此留待未来研究，因为矩阵因子增强模型仍处于起步阶段。

2.2 因子估计量的理论分析

本节研究有限样本情形及因子数过度估计情况下因子估计量的理论性质。我们假设预先设定的常数 m_i, $i = 1, 2$ 不随 p_i 增长，并将其称为“行（列）因子的工作（伪）数量”。由于实践中我们不知道因子的真实数量，通常会选取略大的数值 m_i，使得 $m_i \geq k_i$ 很可能成立。设 $\mathbf{W}_i$ 为确定矩阵或随机矩阵，但与由 $\{\mathbf{E}_t\colon\ t \in [T]\}$ 生成的 σ-代数相互独立。我们进一步假设投影矩阵 $\mathbf{W}_i, i = 1, 2$ 满足以下条件：

假设 2.1. 存在正数常量 c_1 和 c_2，使得（若 $\mathbf{W}_1$,$\mathbf{W}_2$ 为随机矩阵，则几乎必然）当 $p_1, p_2 \to \infty$ 时，

(1) $\max_{i \leq p_1} |w_{1,ij}| < c_1$, $\max_{i \leq p_2} |w_{2,ij}| < c_2$; (2) $m_1 \times m_1$ 矩阵 $\mathbf{W}_1^\top \mathbf{W}_1/p_1$ 和 $m_2 \times m_2$ 矩阵 $\mathbf{W}_2^\top \mathbf{W}_2/p_2$ 满足 $\lambda_{\min}(\mathbf{W}_1^\top \mathbf{W}_1/p_1) \geq c_1$,$\lambda_{\min}(\mathbf{W}_2^\top \mathbf{W}_2/p_2) \geq c_2$;

(3) $\mathbf{W}_1, \mathbf{W}_2$ 与 $\mathbf{E}_t, t \in [T]$ 相互独立。

将矩阵 $\mathbf{F}_t$ 在（2.2）中向量化，我们得到

$$\mathrm{Vec}(\widehat{\mathbf{F}}_t) = (\mathbf{H}_2 \otimes \mathbf{H}_1)\mathrm{Vec}(\mathbf{F}_t) + \frac{1}{p_1 p_2}(\mathbf{W}_2 \otimes \mathbf{W}_1)^\top \mathrm{Vec}(\mathbf{E}_t) := \mathbf{H}\mathrm{Vec}(\mathbf{F}_t) + \frac{1}{p_1 p_2}\mathbf{W}^\top \mathrm{Vec}(\mathbf{E}_t),$$

其中 $\mathbf{H} = \mathbf{H}_2 \otimes \mathbf{H}_1$，$\mathbf{W} = \mathbf{W}_2 \otimes \mathbf{W}_1$。因此，$\mathrm{Vec}(\mathbf{F}_t)$ 可视为 $\mathrm{Vec}(\mathbf{F}_t)$ 的一个估计，二者相差一个变换矩阵，其中 $\mathbf{H} \in \mathbb{R}^{m_1 m_2 \times k_1 k_2}$，估计误差为 $\mathbf{W}^\top \mathrm{Vec}(\mathbf{E}_t)/(p_1 p_2)$。当 $\mathbf{E}_t$ 的每个元素均为截面弱相关时，假设 2.1 保证 $\mathbf{W}^\top \mathrm{Vec}(\mathbf{E}_t)/(p_1 p_2)$ 满足截面中心极限定理。例如，假设 $\mathbf{E}_t$ 的每个元素均为截面独立，在假设 2.1 下，当 $p_1 p_2 \to \infty$ 时，我们得到

$$\frac{1}{\sqrt{p_1 p_2}}\mathbf{W}^\top \mathrm{Vec}(\mathbf{E}_t) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}), \tag{2.3}$$

其中 $\mathbf{V} = \lim_{p_1 p_2 \to \infty} \mathbf{W}^\top \mathrm{Var}(\mathrm{Vec}(\mathbf{E}_t))\mathbf{W}/(p_1 p_2)$，假设其存在。收敛式（2.3）表明，对于每个 $t \leq T$，$\sqrt{p_1 p_2}$，$(\mathrm{Vec}(\mathbf{F}_t) - \mathbf{H}\mathrm{Vec}(\mathbf{F}_t))$ 都是渐近正态的，无论 T 是否趋于无穷大，$m_1 = k_1, m_2 = k_2$ 还是不趋于无穷大。它仅要求 $p_1 p_2 \to \infty$，这对于分析短矩阵时间序列尤其有用。

此外，因子成分不应被分散掉，这意味着变换矩阵 $\mathbf{H}_i, i = 1, 2$ 需满足以下条件。令 $\nu_{\min}(\mathbf{H}_i)$ $\nu_{\max}(\mathbf{H}_i)$ 分别表示 $\mathbf{H}_i$ 的最小和最大非零

奇异值。

假设 2.2. 假设 $m_1 \geq k_1, m_2 \geq k_2$。几乎必然

(1) 秩(H_1) $= k_1$，秩(H_2) $= k_2$;

(2) 存在常数 c_1 和 c_2，使得

$$\nu^2_{\min}(\mathbf{H}_i) \gg \frac{1}{p_i}, \ \nu_{\max}(\mathbf{H}_i) \leq c_1 \nu_{\min}(\mathbf{H}_i), \quad i = 1, 2.$$

假设 2.2 (1) 要求 W_1 至少有 k_1 列与 R 不正交，且 W_2 至少有 k_2 列与 C 不正交，从而 R 和 C 不会被多元化效应掩盖。它还确保$\widehat{\ }$由 Vec(F_t) 张成的空间渐近等于由 Vec(F_t) 张成的空间。这是所施加的关键假设，但在因子过度估计的情形下并不苛刻（Barigozzi 和 Cho，2020年；Fan 和 Liao，2022）。假设 2.2 (2) 决定了恢复因子张成空间的收敛速率，并确保权重矩阵和载荷矩阵不正交。假设$\nu_{\min}(\mathrm{W}_1^\top \mathrm{R}) = p_1^{\alpha_1}, \nu_{\min}(\mathrm{W}_2^\top \mathrm{C}) = p_2^{\alpha_2}$，则假设 2.2 (2) 意味着 $\alpha_1, \alpha_2 \geq 1/2$。假设 2.1 和假设 2.2 是 Fan 和 Liao（2022）中假设 2.1 和 2.2 向矩阵因子模型的直接推广。

对于矩阵因子模型，一旦将观测值向量化，该模型就简化为向量因子模型

$$\mathbf{X} = \mathbf{F}(\mathbf{C} \otimes \mathbf{R})^\top + \mathbf{E},$$

其中 $\mathrm{F} = (\mathrm{Vec}(\mathrm{F}_1), \mathrm{Vec}(\mathrm{F}_2) \ldots, \mathrm{Vec}(\mathrm{F}_T))^\top$，$\mathrm{X} = (\mathrm{Vec}(\mathrm{X}_1), \mathrm{Vec}(\mathrm{X}_2) \ldots, \mathrm{Vec}(\mathrm{X}_T))^\top$。$\mathrm{E} = (\mathrm{Vec}(\mathrm{E}_1), \mathrm{Vec}(\mathrm{E}_2), \ldots, \mathrm{Vec}(\mathrm{E}_T))^\top$。根据 (2.2)，我们有

$$\widehat{\mathbf{F}} = \mathbf{F}\mathbf{H}^\top + \boldsymbol{\mathcal{E}},$$

其中 $\widehat{\ }\mathrm{F} = (\mathrm{Vec}(\widehat{\ }\mathrm{F}_1), \mathrm{Vec}(\widehat{\ }\mathrm{F}_2) \ldots, \mathrm{Vec}(\widehat{\ }\mathrm{F}_T))^\top$，$\mathrm{E} = (\mathrm{Vec}(\mathrm{E}_1), \mathrm{Vec}(\mathrm{E}_2) \ldots, \mathrm{Vec}(\mathrm{E}_T))^\top$。为推导因子估计量的理论性质，我们进一步提出以下假设。

假设 2.3。有常数 c, C 使得

(1) $p_2^{-1}\|\mathbb{E}(\mathrm{E}_t\mathrm{E}_t^\top)\|_2 \leq c$; (2) $1/T \sum_{s=1}^T \sum_{T=1}^T \mathbb{E}\|\mathrm{F}_t\|_F \|\mathrm{F}_s\|_F \mathrm{tr}(\mathbb{E}(\mathrm{E}_t\mathrm{E}_s^\top|\mathrm{F})) < c$; (3) $\lambda_{\min}\left(\mathbb{E}\left(\mathrm{E}^\top\mathrm{E}/T\right)\right) \geq c_0$, $\mathbb{E}\|\mathbb{E}\left(\mathrm{Vec}(\mathrm{E}_t)\mathrm{Vec}(\mathrm{E}_t)^\top|\mathrm{F}\right)\|_2 < c$; (4) $c < \lambda_{\min}(\sum_{t=1}^T \mathrm{Vec}(\mathrm{F}_t)\mathrm{Vec}(\mathrm{F}_t)^\top/T) \leq \lambda_{\max}(\sum_{t=1}^T \mathrm{Vec}(\mathrm{F}_t)\mathrm{Vec}(\mathrm{F}_t)^\top/T) < C$，几乎处处成立; (5) 对任意 $t \in [T], i_1, i_1 \in [p_1], j_1, j_2 \in [p_2]$, $\sum_{s=1}^T \sum_{u_1,v_1=1}^{p_1} \sum_{u_2,v_2=1}^{p_2}|$, $\mathrm{Cov}(e_{t,i_1i_2}e_{t,j_1j_2}, e_{s,u_1u_2}\ e_{s,v_1v_2})| \leq c$。

假设 2.3 要求特质误差之间具有弱依赖性，这在文献中很常见；例如，参见 He 等人 (2021b)、Chen 和 Fan (2021) 以及 Yu 等人 (2022)。以下定理$\widehat{\ }$证明了 $\mathbf{F}_t$ 收敛至真实的 $\mathbf{F}_t$（经过变换矩阵变换后），无论 $T \to \infty$

或高估因子数。

定理 2.1。 假设 2.1 (1)、2.1 (3)、2.2 和 2.3 (1) 均成立。当 $\min\{p_1, p_2\} \to \infty$ 且 T 为有限值或 $T \to \infty$。那么，对于任意有界的 $m_1 \geq k_1$， $m_2 \geq k_2$，对于任意 $t \in [T]$，

$$\left\|\mathbf{M}_1^\top \widehat{\mathbf{F}}_t \mathbf{M}_2 - \mathbf{F}_t\right\|_2 = O_p\left(\frac{1}{\sqrt{p_1}} \nu_{\min}^{-1}(\mathbf{H}_1)\nu_{\min}^{-1}(\mathbf{H}_2)\right),$$

其中 $\mathbf{M}_1 = (\mathbf{H}_1\mathbf{H}_1^\top)^+\mathbf{H}_1 \in \mathbb{R}^{m_1 \times k_1}$， $\mathbf{M}_2 = (\mathbf{H}_2\mathbf{H}_2^\top)^+\mathbf{H}_2 \in \mathbb{R}^{m_2 \times k_2}$。

$\widehat{}$下面的定理表明，由 $\mathbf{F}$ 张成的线性空间渐近地等于由 $\mathbf{F}$ 张成的线性空间。

定理 2.2。 假设 2.1-2.3 成立。对于任意有界的 $m_1 \geq k_1$， $m_2 \geq k_2$，当 $\min\{p_1, p_2\} \to \infty$ 时，我们有

$$\|\mathbf{P}_{\widehat{\mathbf{F}}}\mathbf{P}_{\mathbf{F}} - \mathbf{P}_{\mathbf{F}}\|_2 = O_p\left(\frac{1}{\sqrt{p_1 p_2}} \nu_{\min}^{-1}(\mathbf{H}_1)\nu_{\min}^{-1}(\mathbf{H}_2)\right),$$

$$\|\mathbf{P}_{\widehat{\mathbf{F}}\mathbf{M}} - \mathbf{P}_{\mathbf{F}}\|_2 = O_p\left(\frac{1}{\sqrt{p_1 p_2}} \nu_{\min}^{-1}(\mathbf{H}_1)\nu_{\min}^{-1}(\mathbf{H}_2)\right),$$

其中 $\mathbf{M} = (\mathbf{H}\mathbf{H}^\top)^+\mathbf{H}$，且 $\mathbf{H} = \mathbf{H}_2 \otimes \mathbf{H}_1$。

$\widehat{}$定理 2.2 表明，当 $m_1 \geq k_1$， $m_2 \geq k_2$时，由 $\mathbf{F}$ 张成的线性空间渐近地$\widehat{}$等于由 $\mathbf{F}$张成的线性空间。定理 2.2 还表明，span(**F**) 的一个特定子空间对于 span(**F**) 是一致的。特别地，当 $m_1 = k_1$， $m_2 = k_2$时，我们有 $\mathbf{P}_{\widehat{\mathbf{F}}\mathbf{M}} = \mathbf{P}_{\widehat{\mathbf{F}}}$，因为 $\mathbf{M}$ 可逆。于是它退化为通常的空间一致性。

一致估计因子数量 k_1， k_2 通常需要满足较强的条件，而这些条件在有限样本情况下难以实现。所提出的因子矩阵估计方法的一项优势是，在许多统计推断问题中，即使高估因子数量，该方法仍具有稳健性。关于权重矩阵的选择，可以按照$\mathbf{W}_1$ 和$\mathbf{W}_2$ 的相同策略$\widehat{}$进行选择，Fan 和 Liao (2022)提出了此类策略，例如 Hadamard 投影。根据定理2.2， $\{\mathbf{F}_t, \ t = 1, \ldots\ldots, \ T\}$只要满足 $m_i \geq k_i, \ i = 1, 2$，便可在因子增强模型中得到有效推断。矩阵因子增强模型尚处于起步阶段，因此我们将此留待未来研究。

3 载荷空间的迭代最小二乘估计量

本节介绍列/行载荷空间的迭代最小二乘估计量。我们还推导了在因子数对正确设定的前提下，载荷矩阵估计量在第 s 次迭代时（对任意 $s \geq 1$）的收敛速率。若因子数对未知，我们提出一种新颖的特征值比率法来估计 k_1 和 k_2。

3.1 载荷空间的估计

在第 2 节中，我们介绍了如何利用两个多样化投影矩阵估计因子矩阵$\mathbf{W}_1, \mathbf{W}_2$，并将式（2.2）中的估计量记为 $\{\widehat{\mathbf{F}}_t^{(1)}、\ t = 1、\ldots\ldots T\}$，其维度为 $m_1 \times m_2$，其中$m_1 \geq k_1$，$m_2 \geq k_2$。给定 $\{\widehat{\mathbf{F}}_t^{(1)}\}$，通过在可识别性条件下最小化平方弗罗贝尼乌斯损失，可以直接估计行因子载荷矩阵 $\mathbf{R}$ ：

$$\begin{aligned} &\min_{\mathbf{R}} L_1(\mathbf{R}) = \min_{\mathbf{R}} \frac{1}{T}\sum_{t=1}^{T} \|\mathbf{X}_t - \mathbf{R}\widehat{\mathbf{F}}_t^{(1)}\mathbf{W}_2^\top\|_F^2, \\ &\text{s.t} \quad \frac{1}{p_1}\mathbf{R}^\top\mathbf{R} = \mathbf{I}_{m_1}, m_1 \geq k_1, \end{aligned} \tag{3.1}$$

其中 W　2 是列投影矩阵。式（3.1）中的目标函数可简化为

$$L_1(\mathbf{R}) = \frac{1}{T}\sum_{t=1}^{T}\Big[\mathrm{tr}\big(\mathbf{X}_t^\top\mathbf{X}_t\big) - 2\mathrm{tr}\big(\mathbf{X}_t^\top\mathbf{R}\widehat{\mathbf{F}}_t^{(1)}\mathbf{W}_2^\top\big) + p_1\mathrm{tr}\big(\mathbf{W}_2\widehat{\mathbf{F}}_t^{(1)\top}\widehat{\mathbf{F}}_t^{(1)}\mathbf{W}_2^\top\big)\Big],$$

拉格朗日函数为

$$\min_{\mathbf{R},\boldsymbol{\Theta}} \mathcal{L}_1(\mathbf{R},\boldsymbol{\Theta}) = \min_{\mathbf{R},\boldsymbol{\Theta}} \left\{ L_1(\mathbf{R}) + \mathrm{tr}\left[\boldsymbol{\Theta}\left(\frac{1}{p_1}\mathbf{R}^\top\mathbf{R} - \mathbf{I}_{m_1}\right)\right]\right\},$$

其中，拉格朗日乘子 $\boldsymbol{\Theta}$ 是一个对称矩阵。令 $\partial\mathcal{L}_1(\mathrm{R}\ \boldsymbol{\Theta})/\partial\mathrm{R} = 0$ 和 $\partial\mathcal{L}_1(\mathrm{R}\ \boldsymbol{\Theta})/\partial\boldsymbol{\Theta} = 0$，可得

$$\frac{\partial\mathcal{L}_1(\mathbf{R},\boldsymbol{\Theta})}{\partial\mathbf{R}} = \frac{1}{T}\sum_{t=1}^{T}\big(-2\mathbf{X}_t\mathbf{W}_2\widehat{\mathbf{F}}_t^{(1)\top} + \frac{2}{p_1}\mathbf{R}\boldsymbol{\Theta}\big) = 0, \quad \frac{\partial\mathcal{L}_1(\mathbf{R},\boldsymbol{\Theta})}{\partial\boldsymbol{\Theta}} = \frac{1}{p_1}\mathbf{R}^\top\mathbf{R} - \mathbf{I}_{m_1} = 0. \tag{3.2}$$

此外，我们可以推导出 $\boldsymbol{\Theta}$ 和 $\mathbf{R}$ 满足（3.2）的显式表达式，即:

$$\widehat{\boldsymbol{\Theta}} = \frac{\sqrt{p_1}}{T}\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^{(1)}\mathbf{W}_2^\top\mathbf{X}_t^\top\right)\left(\sum_{t=1}^{T}\mathbf{X}_t\mathbf{W}_2\widehat{\mathbf{F}}_t^{(1)\top}\right)\right]^{1/2},$$

$$\widehat{\mathbf{R}}^{(1)} = \sqrt{p_1}\left(\sum_{t=1}^{T}\mathbf{X}_t\mathbf{W}_2\widehat{\mathbf{F}}_t^{(1)\top}\right)\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^{(1)}\mathbf{W}_2^\top\mathbf{X}_t^\top\right)\left(\sum_{t=1}^{T}\mathbf{X}_t\mathbf{W}_2\widehat{\mathbf{F}}_t^{(1)\top}\right)\right]^{-1/2}, \tag{3.3}$$

并将 $\widehat{\mathbf{R}}^{(1)}$ 设为行载荷矩阵的一步估计量。类似地，一旦得到估计量$\widehat{\mathbf{R}}^{(1)}$，我们便可在可识别性条件下，通过最小化以下损失函数得到 $\widehat{\mathbf{C}}^{(1)}$：

$$\begin{aligned}&\min_{\mathbf{C}} L_2(\mathbf{C}) = \min_{\mathbf{C}} \frac{1}{T}\sum_{t=1}^{T} \left\|\mathbf{X}_t - \widehat{\mathbf{R}}^{(1)}\widehat{\mathbf{F}}_t^{(1)}\mathbf{C}^\top\right\|_F^2,\\ &\text{s.t}\frac{1}{p_2}\mathbf{C}^\top\mathbf{C} = \mathbf{I}_{m_2}, m_2 \geq k_2,\end{aligned}$$

利用拉格朗日函数

$$\min_{\mathbf{C},\mathbf{\Lambda}} \mathcal{L}_2(\mathbf{C},\mathbf{\Lambda}) = \min_{\mathbf{C},\mathbf{\Lambda}} \left\{\frac{1}{T}\sum_{t=1}^{T} \left\|\mathbf{X}_t - \widehat{\mathbf{R}}^{(1)}\widehat{\mathbf{F}}_t^{(1)}\mathbf{C}^\top\right\|_F^2 + \operatorname{tr}\left[\mathbf{\Lambda}\left(\frac{1}{p_2}\mathbf{C}^\top\mathbf{C} - \mathbf{I}_{m_2}\right)\right]\right\},$$

我们得到列因子载荷矩阵的以下估计量

$$\widehat{\mathbf{C}}^{(1)} = \sqrt{p_2}\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(1)}\widehat{\mathbf{F}}_t^{(1)}\right)\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^{(1)\top}\widehat{\mathbf{R}}^{(1)\top}\mathbf{X}_t\right)\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(1)}\widehat{\mathbf{F}}_t^{(1)}\right)\right]^{-1/2}. \tag{3.4}$$

给定 $\widehat{\mathbf{C}}^{(1)}$，我们将 $\mathbf{F}_t$ 的估计量更新为 $\widehat{\mathbf{F}}_t^{(2)} = \widehat{\mathbf{R}}^{(1)\top}\mathbf{X}_t\widehat{\mathbf{C}}^{(1)}/(p_1p_2)$，从而更新 R 和 C 的估计量。具体而言，在 (3.3) 中用 $\widehat{\mathbf{F}}_t^{(2)}$、$\widehat{\mathbf{C}}^{(1)}$ 分别替换 $\widehat{\mathbf{F}}_t^{(1)}$、$\mathbf{W}_2$，以此更新 R 的估计量 $\widehat{\mathbf{R}}^{(2)}$；在 (3.4) 中用 $\widehat{\mathbf{F}}_t^{(2)}$、$\widehat{\mathbf{R}}^{(2)}$ 分别替换 $\widehat{\mathbf{F}}_t^{(1)}$、$\widehat{\mathbf{R}}^{(1)}$，以此更新 C 的估计量 $\widehat{\mathbf{C}}^{(2)}$；重复上述步骤，直至满足收敛准则。在第 $(s+1)$ 次迭代中，估计量$\widehat{\mathbf{F}}_t^{(s+1)}$、$\widehat{\mathbf{R}}^{(s+1)}$ 和 $\widehat{\mathbf{C}}^{(s+1)}$ 的表达式如下：

$$\widehat{\mathbf{F}}_t^{(s+1)} = \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(s)\top}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)},$$

$$\widehat{\mathbf{R}}^{(s+1)} = \sqrt{p_1}\left(\sum_{t=1}^{T}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{F}}_t^{(s+1)\top}\right)\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^{(s+1)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{X}_t^\top\right)\left(\sum_{t=1}^{T}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{F}}_t^{(s+1)\top}\right)\right]^{-1/2},$$

$$\widehat{\mathbf{C}}^{(s+1)} = \sqrt{p_2}\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{F}}_t^{(s+1)}\right)\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^{(s+1)\top}\widehat{\mathbf{R}}^{(s+1)\top}\mathbf{X}_t\right)\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{F}}_t^{(s+1)}\right)\right]^{-1/2},$$

其中 $\widehat{\mathbf{R}}^{(s)}$ 和 $\widehat{\mathbf{C}}^{(s)}$ 是第 s 次迭代得到的估计量。基于随机投影的迭代最小二乘（RPILS）矩阵因子模型过程总结于下文的算法 1 中，理论分析将在下一节给出。

对于收敛准则，将第（$s+1$）步的共同成分矩阵记为 $\widehat{\mathbf{S}}_t^{(s+1)} = \widehat{\mathbf{R}}^{(s)}\widehat{\mathbf{F}}_t^{(s+1)}\widehat{\mathbf{C}}^{(s)\top}$。在我们的模拟研究中，迭代过程会在达到预先设定的

最大迭代次数（maxiter = 100）或当

$$\|\widehat{\mathbf{S}}_t^{(s+1)} - \widehat{\mathbf{S}}_t^{(s)}\|_F \leq \epsilon,$$

其中 ϵ 是预先给定的一个小常数（10^{-6}）。

算法1 基于随机投影的迭代最小二乘（RPILS）矩阵因子模型程序

输入：数据矩阵 $\{\mathbf{X}_t\}$、$t \leq T$，伪行列因子数对 m_1（$m_1 \geq k_1$）以及 m_2（$m_2 \geq k_2$）、多样化投影矩阵 $\mathbf{W}_1$,$\mathbf{W}_2$

输出：因子载荷矩阵 $\widetilde{\mathbf{R}} \in \mathbb{R}^{p_1 \times m_1}$, $\widetilde{\mathbf{C}} \in \mathbb{R}^{p_2 \times m_2}$ 和因子矩阵 $\widetilde{\mathbf{F}}_t \in \mathbb{R}^{m_1 \times m_2}, t \leq T$

1: 通过 $\widehat{\mathbf{F}}_t^{(1)} = \mathbf{W}_1^\top \mathbf{X}_t \mathbf{W}_2/(p_1 p_2)$，得到初始估计量 $\widehat{\mathbf{F}}_t^{(1)}$；

2: 给定 $\widehat{\mathbf{F}}_t^{(1)}$ 和 $\mathbf{W}_2$, , 通过方程（$\widehat{\mathbf{R}}^{(1)}$ 3.3）得到估计量；给定 $\widehat{\mathbf{F}}_t^{(1)}$ 和 $\widehat{\mathbf{R}}^{(1)}$，得到估计量 $\widehat{\mathbf{C}}^{(1)}$ 通过方程（3.4）；

3: 通过 $\widehat{\mathbf{F}}_t^{(2)} = \widehat{\mathbf{R}}^{(1)\top} \mathbf{X}_t \widehat{\mathbf{C}}^{(1)}/(p_1 p_2)$更新 $\widehat{\mathbf{F}}_t^{(2)}$；

4: 通过将 $\widehat{\mathbf{R}}^{(2)}$ 替换为 $\widehat{\mathbf{F}}_t^{(1)}$、$\mathbf{W}_2$，将方程（3.3）中的 $\widehat{\mathbf{F}}_t^{(2)}$、$\widehat{\mathbf{C}}^{(1)}$；分别更新；通过 $\widehat{\mathbf{C}}^{(2)}$ 替换将 $\widehat{\mathbf{F}}_t^{(1)}$、$\widehat{\mathbf{R}}^{(1)}$ 替换为方程（3.4）中的 $\widehat{\mathbf{F}}_t^{(2)}$、$\widehat{\mathbf{R}}^{(2)}$，分别替换；

5: 重复步骤 3-4，直至收敛，并输出最后一步得到的估计量，记为 $\widetilde{\mathbf{R}}$, $\widetilde{\mathbf{C}}$ 和 $\{\widetilde{\mathbf{F}}_t, t \leq T\}$，分别。

3.2 理论性质

本节将建立任意 $s \geq 1$ 下第 s 次迭代时载荷矩阵估计量的收敛速率。为此，我们首先提出一些文献中常见的附加条件。

假设3.1。因子矩阵满足 $\mathbb{E}(\mathbf{F}_t) = \mathbf{0}, \mathbb{E}\|\mathbf{F}_t\|^4 \leq c < \infty$，其中 $c > 0$ 为某个常数，且

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{F}_t \mathbf{F}_t^\top \xrightarrow{p} \mathbf{\Sigma}_1 \quad \text{and} \quad \frac{1}{T}\sum_{t=1}^{T} \mathbf{F}_t^\top \mathbf{F}_t \xrightarrow{p} \mathbf{\Sigma}_2,$$

其中 $\mathbf{\Sigma}_i, i = 1, 2$ 是特征值有界的正定矩阵。

假设 3.2。存在正数常量 $c, c_0 < \infty$，使得 (1) $\mathbb{E}(e_{t,ij}) = 0$. (2) 对任意 $t \in [T], i \in [p_1], j \in [p_2]$, $\sum_{s=1}^{T}\sum_{l=1}^{p_1}\sum_{h=1}^{p_2} |\mathbb{E}(e_{t,ij}e_{s,lh})| \leq c$ 以及 $\sum_{l=1}^{p_1}\sum_{h=1}^{p_2} |\mathbb{E}(e_{t,lj}e_{t,ih})| \leq c$.(3) 对任意 $T \in [T], i_1, i_1 \in [p_1], j_1, j_2 \in [p_2]$, $\sum_{s=1}^{T}\sum_{u_1,v_1=1}^{p_1}\sum_{u_2,v_2=1}^{p_2}$, $\mathrm{Cov}(e_{t,i_1 i_2}e_{t,j_1 j_2}, e_{s,u_1 u_2}e_{s,v_1 v_2})| \leq c$。

Assu假设 3.3。存在正数常量 c_1、c_2，使得 $\|\mathbf{R}\|_{\max} \leq c_1$、$\|\mathbf{C}\|_{\max} \leq c_2$.

假设 3.4。存在常数 $c > 0$，使得：(1) 对于满足 $\|\boldsymbol{v}\| = 1$ 和 $\|\boldsymbol{w}\| = 1$ 的任意确定性向量 $\boldsymbol{v}$ 和 $\boldsymbol{w}$，有 $\mathbb{E}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\mathbf{F}_t \boldsymbol{v}^\top \mathbf{E}_t \boldsymbol{w})\right\|^2 \leq c$；(2) 对于任意 $i \in [p_1]$、$j \in [p_2]$、$\|\sum_{i'=1}^{p_1}\sum_{j'=1}^{p_2}\mathbb{E}(\boldsymbol{\xi}_{i,\ j} \otimes \boldsymbol{\xi}_{i',\ j'})\|_{\max} \leq c$，其中 $\boldsymbol{\xi}_{i,\ j} = \mathrm{Vec}(T^{-1/2}\sum_{t=1}^{T}\mathbf{F}_t e_{t,\ ij})$；(3) 对于任意 i_1、$i_2 \in [p_1]$、j_1、$j_2 \in [p_2]$、$\|\sum_{i_1'}^{p_1}$、$\sum_{i_1'=1}^{p_2}$、$\sum_{j_2'=1}^{j_1'}$、$\mathrm{Cov}(\boldsymbol{\xi}_{i_1,\ j_1} \otimes \boldsymbol{\xi}_{i_2,\ j_2}, \boldsymbol{\xi}_{i_1',\ j_1'} \otimes \boldsymbol{\xi}_{i_2',\ j_2'})\|_{\max} \leq c$.

假设 3.1–3.4 是矩阵因子模型文献中的标准假设；例如参见 Chen 和 Fan（2021）、Yu 等（2022）以及 He 等（2021b）。假设 3.2 确保特质项 E_t 在（截面和时间序列意义下）可求和，从而允许空间和时间域中存在（弱）依赖。假设3.3 要求公共因子具有普遍性。假设 3.4 允许公共因子 F_t 与误差 E_t 弱相关；例如，当 $\{\mathrm{F}_t\}$ 和 $\{\mathrm{E}_t\}$ 是两个相互独立的组时，该假设便得到满足。关于这些假设的详细讨论，可参阅 He 等（2021b）。在下面的定理中，我们首先建立分别在（3.3）和（3.4）中定义的一步估计量 $\widehat{\mathrm{R}}^{(1)}$ 和 $\widehat{\mathrm{C}}^{(1)}$ 的收敛速率。我们假设本节中的 $m_1 = k_1$、$m_2 = k_2$ 以及因子数量 k_1、k_2 均为固定值。在第 3.3 节中，我们提出一种估计因子数量的方法。

定理 3.1。 在假设 2.1 (1), 2.1 (3), 2.2，假设 3.1-3.4成立的条件下，进一步假设 $\nu^2_{\min}(\mathbf{H}_2) \gg \max(1/T, 1/p_2)$，当 T, p_1, p_2 趋于无穷大时，存在一个渐近正交矩阵 $\widehat{\mathbf{H}}^{(1)}_r$，使得

$$\frac{1}{p_1}\|\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}^{(1)}_r\|^2_F = O_p\left(\frac{1}{Tp_2\nu^2_{\min}(\mathbf{H}_2)} + \frac{1}{Tp_1p_2\nu^2_{\min}(\mathbf{H}_1)\nu^4_{\min}(\mathbf{H}_2)} + \frac{1}{p_1^2p_2^2\nu^2_{\min}(\mathbf{H}_1)\nu^4_{\min}(\mathbf{H}_2)}\right), \quad (3.5)$$

其中 $\mathrm{H}_1 = \mathrm{W}_1^\top \mathrm{R}/p_2$，$\mathrm{H}_2 = \mathrm{W}_2^\top \mathrm{C}/p_2$。

条件 $\nu^2_{\min}(\mathrm{H}_2) \gg \max(1/T, 1/p_2)$ 保证矩阵 $\widehat{\mathrm{H}}_r$ 渐近正交。实质上，该条件要求初始投影矩阵 W_2 的列所张成的空间与 C 的列所张成的空间相差不大；如果这两个空间正交，则该条件不成立。可以想见，如果从真实空间的正交空间开始，迭代算法将永远无法收敛到真实空间。定理 3.1 表明，初始投影方向（即 W_2 的列所张成的空间）越接近真实载荷方向（即 C 的列所张成的空间），估计载荷矩阵 $\widehat{\mathrm{R}}^{(1)}$ 收敛到真实载荷矩阵 R（相差一个正交矩阵）的速度就越快。特别地，如果 $\nu_{\min}(\mathrm{H}_1) = \nu_{\min}(\mathrm{H}_2) = O_p(1)$（只要我们以 Chen 和 Fan（2021）中一致的 α-主成分分析估计量作为投影方向），则有

$$\frac{1}{p_1}\|\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}^{(1)}_r\|^2_F = O_p\left(\frac{1}{Tp_2} + \frac{1}{p_1^2p_2^2}\right).$$

在下述定理中，我们给出一步估计量的收敛速率 $\widehat{\mathbf{C}}^{(1)}$ 定义于（3.4）。

定理 3.2。 在与定理 3.1相同的条件下，存在一个渐近正交矩阵$\widehat{\mathbf{H}}^{(1)}_c$，使得

$$w_c^{(1)} = \frac{1}{p_2}\|\widehat{\mathbf{C}}^{(1)} - \mathbf{C}\widehat{\mathbf{H}}^{(1)}_c\|^2_F = O_p\left(\frac{1}{Tp_1} + \frac{1}{p_1^2p_2^2\nu^2_{\min}(\mathbf{H}_1)\nu^2_{\min}(\mathbf{H}_2)} + \frac{1}{T^2p_2^2\nu^2_{\min}(\mathbf{H}_2)} + \frac{1}{Tp_1p_2\nu^2_{\min}(\mathbf{H}_1)\nu^2_{\min}(\mathbf{H}_2)}\right).$$

此外，$w_r^{(1)} = \frac{1}{p_1}\|\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\,\widehat{\mathbf{H}}_r^{(1)}\|_F^2$是式（3.5）中推导出的速率。

假设 $\nu_{\min}(\mathbf{H}_1) = \nu_{\min}(\mathbf{H}_2) = O_p(1)$，则推导出的 $\widehat{\mathbf{C}}^{(1)}$ 收敛速率为

$$\frac{1}{p_2}\|\widehat{\mathbf{C}}^{(1)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(1)}\|_F^2 = O_p\left(\frac{1}{Tp_1} + \frac{1}{p_1^2p_2^2} + \frac{1}{T^2p_2^2}\right),$$

这与Yu 等人(2022)中推论 3.1 推导出的结果相同。

只要得到行/列载荷矩阵的估计量，即 $\widehat{\mathbf{R}}^{(1)}$ 和 $\widehat{\mathbf{C}}^{(1)}$，就可以通过 $\widehat{\mathbf{F}}_t^{(2)} = \widehat{\mathbf{R}}^{(1)\top}\mathbf{X}_t\widehat{\mathbf{C}}^{(1)}/(p_1p_2)$ 更新因子矩阵，随后将相应的共同成分矩阵更新为 $\widehat{\mathbf{S}}_t^{(2)} = \widehat{\mathbf{R}}^{(1)}\widehat{\mathbf{F}}_t^{(2)}\widehat{\mathbf{C}}^{(1)\top} = (\widehat{S}_{t,\,ij})$。以下定理给出经过一次迭代后估计因子和共同成分的收敛速率。

定理 3.3。在与定理3.1相同的条件下，当 $\min\{T, p_1, p_2\} \to \infty$时，对于任意 $t \in [T]$，$i \in [p_1]$ 和 $j \in [p_2]$，有

$$\left\|\widehat{\mathbf{F}}_t^{(2)} - (\widehat{\mathbf{H}}_r^{(1)})^{-1}\mathbf{F}_t\left((\widehat{\mathbf{H}}_c^{(1)})^{-1}\right)^\top\right\|_F = O_p\left(\frac{1}{\sqrt{T}p_1} + \frac{1}{\sqrt{T}p_2\nu_{\min}(\mathbf{H}_2)} + \frac{1}{\sqrt{p_1p_2}} + \gamma_f\right),$$

其中

$$\gamma_f = \frac{1}{\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)} + \frac{1}{p_1p_2\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)},$$

对于公共成分，我们有

$$|\widehat{S}_{t,ij}^{(2)} - S_{t,ij}| = O_p\left(\frac{1}{\sqrt{Tp_1}} + \frac{1}{\sqrt{p_1p_2}} + \frac{1}{\sqrt{Tp_2}\nu_{\min}(\mathbf{H}_2)} + \frac{1}{\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)} + \frac{1}{p_1p_2\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}\right),$$

其中 $S_{t,\,ij}$ 是 $\mathbf{S}_t = \mathbf{R}\mathbf{F}_t\mathbf{C}^\top$ 的第（i，j）个元素。

定理 3.3 中推导出的收敛速率与 Yu 等人（2022）在 $\nu_{\min}(\mathbf{H}_1) = \nu_{\min}(\mathbf{H}_2) = O_p(1)$ 中的定理 3.5 推导出的收敛速率相同。在下面的定理中，对于任意整数 $\widehat{\mathbf{R}}^{(s+1)}$ 和 $\widehat{\mathbf{C}}^{(s+1)}$，我们建立估计量 $s \geq 1$ 的收敛速率递推公式。

定理 3.4。在定理 3.1所述的相同条件下，存在渐近正交矩阵$\widehat{\mathbf{H}}_r^{(s+1)}$和 $\widehat{\mathbf{H}}_c^{(s+1)}$，使得对于任意整数 $s \geq 1$，有

$$w_r^{(s+1)} = \frac{1}{p_1}\|\widehat{\mathbf{R}}^{(s+1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}\|_F^2 = O_p\left(\frac{1}{Tp_2} + \frac{1}{p_1^2p_2^2} + \gamma_r^{(s+1)}\right),$$

$$w_c^{(s+1)} = \frac{1}{p_2}\|\widehat{\mathbf{C}}^{(s+1)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(s+1)}\|_F^2 = O_p\left(\frac{1}{Tp_1} + \frac{1}{p_1^2p_2^2} + \gamma_c^{(s+1)}\right),$$

其中

$$\gamma_r^{(s+1)} = \frac{1}{p_2^2}w_r^{(s)} + \frac{1}{T}w_c^{(s)} + \frac{w_r^{(s)}w_c^{(s)}}{p_2} + w_r^{(s)}w_c^{(s)2} + \frac{1}{p_1^2}w_c^{(s)2},$$

$$\gamma_c^{(s+1)} = \frac{1}{Tp_2} w_r^{(s)} + \frac{1}{T} w_r^{(s+1)} + \frac{1}{p_1^2} w_c^{(s)} + \frac{1}{p_1} w_r^{(s)} w_c^{(s)} + \frac{1}{T} w_r^{(s)} w_c^{(s)} + w_r^{(s)} w_r^{(s+1)} w_c^{(s)} + \frac{1}{p_2^2} w_r^{(s)} w_r^{(s+1)}.$$

在以下定理中，我们还建立了估计量 $\widehat{\mathbf{F}}^{(s+1)}$ 和 $\widehat{\mathbf{S}}_t^{(s+1)}$ 在给定任意整数 $s \geq 1$ 时的收敛速率递推公式。

定理 3.5。 在定理 3.1所述的相同条件下，当 $\min\{T, p_1, p_2\} \to \infty$时，对任意 $t \in [T]$，$i \in [p_1]$ 以及 $j \in [p_2]$ 和任意整数 $s > 1$，有

$$\left\| \widehat{\mathbf{F}}_t^{(s+1)} - (\widehat{\mathbf{H}}_r^{(s)})^{-1} \mathbf{F}_t \left((\widehat{\mathbf{H}}_c^{(s)})^{-1} \right)^\top \right\|_F = O_p \left(\sqrt{\frac{w_r^{(s)}}{p_2}} + \sqrt{\frac{w_r^{(s-1)}}{Tp_2}} + \sqrt{\frac{w_c^{(s)}}{p_1}} + \sqrt{\frac{w_c^{(s-1)}}{Tp_1}} + \frac{1}{\sqrt{p_1 p_2}} + \gamma_f^{(s+1)} \right),$$

其中

$$\gamma_f^{(s+1)} = \sqrt{w_r^{(s)} w_c^{(s)}} + \sqrt{\frac{w_r^{(s-1)} w_c^{(s-1)}}{T}} + \frac{\sqrt{w_r^{(s-1)} w_c^{(s-1)}}}{p_1},$$

对于公共成分，我们有

$$|\widehat{S}_{t,ij}^{(s+1)} - S_{t,ij}| = O_p \left(\sqrt{\frac{w_r^{(s)}}{p_2}} + \sqrt{\frac{w_r^{(s-1)}}{Tp_2}} + \sqrt{\frac{w_c^{(s)}}{p_1}} + \sqrt{\frac{w_c^{(s-1)}}{Tp_1}} + \frac{1}{\sqrt{p_1 p_2}} + \gamma^{(s+1)} \right),$$

其中 $\gamma^{(s+1)} = \sqrt{\frac{w_r^{(s-1)} w_c^{(s-1)}}{T}} + \frac{\sqrt{w_r^{(s-1)} w_c^{(s-1)}}}{p_1}$和 $S_{t,\ ij}$ 是 $\mathbf{S}_t = \mathbf{R}\mathbf{F}_t\mathbf{C}^\top$ 的第（i，j）个元素。

上述载荷矩阵估计量的理论分析依赖于因子数的正确设定（注意，我们假设 $m_1 = k_1$，$m_2 = k_2$ 在定理 3.1中）以及强因子条件$\mathbf{R}^\top\mathbf{R}/p_1 = \mathbf{I}_{k_1}$，$\mathbf{C}^\top\mathbf{C}/p_2 = \mathbf{I}_{k_2}$，这意味着行因子和列因子在两个维度上都具有广泛性，是对 Stock 和 Watson (2002a) 中广泛性假设在矩阵情形下的扩展。

3.3 确定因子数量

在实际应用中，式（1.1）中的公共因子矩阵 $\mathbf{F}_t$ 的维度 k_1 和 k_2 未知，需要加以确定。本研究借鉴 Lam and Yao (2012) 和 Ahn and Horenstein (2013) 中讨论的特征值比率思想，确定行因子和列因子的数量。不过，我们的方法基于因子数较大的一对伪值所对应的估计因子矩阵的行/列样本协方差矩阵（的特征值比率）；与此不同，文献中的方法（如 Wang 等（2019）；Chen 和 Fan（2021）；Yu 等（2022））基于原始/投影观测值的行/列样本协方差矩阵。这一点本身也值得关注。具体而言，我们首先设定 $m_1 = m_2 = k_{\max}$（$k_{\max} > \max\{k_1, k_2\}$）和~，由此通过算法1可得到 $\mathbf{F}_t$，其维度为 $k_{\max} \times k_{\max}$。随后进一步按如下方式确定维度 k_1 和 k_2:

$$\widehat{k}_1 = \arg\max_{j\le k_{\max}} \frac{\lambda_j\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top\right)}{\lambda_{j+1}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top\right)}, \quad \widehat{k}_2 = \arg\max_{j\le k_{\max}} \frac{\lambda_j\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{F}}_t^\top\widetilde{\mathbf{F}}_t\right)}{\lambda_{j+1}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{F}}_t^\top\widetilde{\mathbf{F}}_t\right)}. \tag{3.6}$$

我们有 th 以下定理保证了式（3.6 ）.

定理 3.6。 在满足定理 3.1的相同假设下，若 $k_{\max} >$最大值$\{k_1, k_2\}$，当 $T, p_1, p_2 \to \infty$,
$\widehat{\ }$ $\widehat{\ }$则有 $\mathbb{P}(k_1 = k_1) \to 1$ 以及 $\mathbb{P}(k_2 = k_2) \to 1$。

在第 3.2 节的理论分析中，假设因子数事先已知；而在第 2 节中，即使高估因子数量，对因子空间的推断仍具有稳健性。定理 3.6 弥补了这一缺口，因为我们总能以趋近于 1 的概率正确确定因子数。

4 模拟研究

4.1 数据生成

本节中，我们研究基于随机投影的迭代最小二乘（RPILS）方法在估计载荷空间和因子空间方面的实证表现。我们首先介绍合成数据集的数据生成机制，该机制与 He 等人 (2022b) 类似。设 $k_1 = 3, k_2 = 3$，从均匀分布 $\mathbf{R}$ 和 $\mathbf{C}$ 中独立抽取各元素 $U(-1, 1)$，并令

$$\mathbf{F}_t = \phi\mathbf{F}_{t-1} + \sqrt{1-\phi^2}\boldsymbol{\epsilon}_t, \quad \mathbf{E}_t = \psi\mathbf{E}_{t-1} + \sqrt{1-\psi^2}\mathbf{U}_t,$$

其中，$\mathrm{Vec}(\boldsymbol{\epsilon}_t) \overset{i.i.d}{\sim} \mathcal{N}(0, \mathrm{I}_{k_1\times k_2})$，$\mathrm{U}_t \overset{i.i.d}{\sim} \mathcal{MN}(0, \mathrm{U}_E, \mathrm{V}_E)$，即 $\mathrm{Vec}(\mathrm{U}_t) \overset{i.i.d}{\sim} \mathcal{N}(0, \mathrm{V}_E \otimes \mathrm{U}_E)$。参数$\phi$ 和 ψ 控制时间相关性，U_E 和 V_E 是对角线上元素为 1 的矩阵，其非对角线元素分别为 $1/p_1$ 和 $1/p_2$。

4.2 载荷空间的估计误差

我们比较基于随机投影的迭代最小二乘（RPILS）方法与 α-主成分分析方法（$\alpha = 0$），后者由Chen 和 Fan (2021)提出；同时比较 Yu 等人提出的 PE 方法（Yu 等人 (2022)），比较内容是对载荷空间和因子空间的估计。对于 RPILS 方法，初始权重矩阵$\mathbf{W}_1$,$\mathbf{W}_2$采用 Walsh-Hadamard 矩阵（Fan 和 Liao，2022），列维度 m_1、m_2 设为因子矩阵的真实维度，即 $m_1 = k_1 = 3$，$m_2 = k_2 = 3$。为展示初始权重矩阵的影响，我们还将其与单步估计量（OSE）进行比较：OSE 的初始权重矩阵采用 Walsh-Hadamard 矩阵，或采用 α-主成分分析估计量作为初始权重矩阵，分别记为 OSE_1 和 OSE_2

分别。我们指出，在未报告的模拟实验中，我们尝试将元素均来自标准正态分布的矩阵用作 RPILS 方法的初始权重矩阵，发现其性能与采用 Walsh-Hadamard 矩阵作为初始权重矩阵时几乎相同。我们考虑以下两种参数设置情形：

情景A： $p_1 = 20, T = p_2 \in \{20, 50, 100, 150, 200\}, \phi = 0.1, \psi = 0.1.$

情景B： $p_2 = 20, T = p_1 \in \{20, 50, 100, 150, 200\}, \phi = 0.1, \psi = 0.1.$

为了衡量各种方法在估计载荷/因子空间方面的表现，我们采用一种线性空间之间的度量，该度量也被 Yu 等人 (2022); He 等人 (2021a) 使用。对于两个按列正交的矩阵（$\boldsymbol{Q}_1$）$_{p\times q_1}$ 和（$\boldsymbol{Q}_2$）$_{p\times q_2}$，我们定义

$$\mathcal{D}(\boldsymbol{Q}_1, \boldsymbol{Q}_2) = \left(1 - \frac{1}{\max(q_1, q_2)} \mathrm{Tr}\left(\boldsymbol{Q}_1 \boldsymbol{Q}_1^\top \boldsymbol{Q}_2 \boldsymbol{Q}_2^\top\right)\right)^{1/2}.$$

根据 $\mathcal{D}(\boldsymbol{Q}_1, \boldsymbol{Q}_2)$的定义，我们很容易看出 $0 \le \mathcal{D}(\boldsymbol{Q}_1, \boldsymbol{Q}_2) \le 1$，它衡量由 $\boldsymbol{Q}_1$ 和 $\boldsymbol{Q}_2$张成的列空间之间的距离，即 span($\boldsymbol{Q}_1$) 和 span($\boldsymbol{Q}_2$)。特别地，当 $\mathcal{D}(\boldsymbol{Q}_1, \boldsymbol{Q}_2) = 0$时，span($\boldsymbol{Q}_1$) 和 span($\boldsymbol{Q}_2$) 相同；而当$\mathcal{D}(\boldsymbol{Q}_1, \boldsymbol{Q}_2) = 1$时，span($\boldsymbol{Q}_1$) 和 span($\boldsymbol{Q}_2$) 正交。Gram-Schmidt正交化可用于将 $\boldsymbol{Q}_1$ 和 $\boldsymbol{Q}_2$ 变为按列正交的矩阵。

表1 报告了在 $\mathcal{D}(\mathbf{R}, \mathbf{R})$, $\mathcal{D}(\mathbf{C}, \mathbf{C})$和 $\mathcal{D}(\mathrm{Vec}(\mathbf{F}_t), \mathrm{Vec}(\mathbf{F}_t))$方面的平均估计误差（括号内为标准误），基于情景A和B的500次重复试验计算得出。就载荷空间估计而言，所有方法均受益于较大的维度。比较OSE_1 和OSE_2, 的结果后，我们得出结论：初始投影方向越好，载荷/因子空间就越快收敛到相应的真实空间。对于RPILS，结果表明，即使从随机猜测的投影方向开始，通过算法1中的迭代过程，最终也能获得令人满意的估计量。此外，RPILS的表现与PE方法相当，优于 α-主成分分析方法，这与我们的理论分析一致。换句话说，RPILS方法也会降低特异性误差成分的幅度，从而提高信噪比，具有与PE相同的优势。然而，与基于特征分解的PE方法相比，RPILS方法在计算上更简单。

4.3 估计因子数量

在本节中，我们比较所提出的 RPILS-ER 方法与 α-PCA 方法的实证表现（α-PCA-ER），后者由Chen 和 Fan (2021) 提出；此外，还比较了 Yu 等人(2022) 提出的 IterER 方法、基于 iTOPUP 和 iTIPUP 的 ER 方法（Han 等人(2022)）（分别记为 iTOP-ER 和 iTIP-ER），以及

表1：在矩阵正态分布下，针对情景A和B进行500次重复试验所得的 $\mathcal{D}$(R, R)、$\mathcal{D}$(C,C)和 $\mathcal{D}$(Vec(F_t),Vec(F_t))方面的平均估计误差（括号内为标准误）。

评估	T	p_1	p_2	OSE_1	OSE_2	RPILS	α-主成分分析	PE
				设置A： $p_1 = 20,\ \ p_2 = T$				
$\mathcal{D}$(R, R)	20	20	20	0.6178(0.1205)	0.0954(0.0173)	0.0938(0.0158)	0.1151(0.0308)	0.0947(0.0162)
	50	20	50	0.5992(0.1197)	0.0357(0.0053)	0.0355(0.0052)	0.0588(0.0219)	0.0355(0.0052)
	100	20	100	0.5797(0.1330)	0.0177(0.0025)	0.0176(0.0025)	0.0459(0.0221)	0.0176(0.0025)
	150	20	150	0.5835(0.1315)	0.0117(0.0017)	0.0117(0.0016)	0.0426（0.0199）	0.0117（0.0016）
	200	20	200	0.5762（0.1305）	0.0088（0.0013）	0.0088（0.0012）	0.0415（0.0203）	0.0088(0.0012)
$\mathcal{D}$(C, C)	20	20	20	0.6138(0.1182)	0.0947(0.0166)	0.0933(0.0156)	0.1148(0.0290)	0.0942(0.0160)
	50	20	50	0.5458(0.1446)	0.0572(0.0063)	0.0568(0.0062)	0.0594(0.0068)	0.0572(0.0063)
	100	20	100	0.4785(0.1508)	0.0402(0.0034)	0.0399(0.0033)	0.0407(0.0035)	0.0402(0.0034)
	150	20	150	0.4400(0.1643)	0.0326(0.0026)	0.0324(0.0025)	0.0328(0.0025)	0.0326(0.0026)
	200	20	200	0.4217(0.1619)	0.0282(0.0021)	0.0280(0.0021)	0.0282(0.0021)	0.0282(0.0021)
$\mathcal{D}$(Vec(F_t), Vec(F_t))	20	20	20	0.6740(0.0380)	0.1837(0.0371)	0.1783(0.0338)	0.1837(0.0371)	0.1784(0.0339)
	50	20	50	0.7928(0.0440)	0.1074(0.0122)	0.1059(0.0111)	0.1074(0.0122)	0.1059(0.0111)
	100	20	100	0.8213(0.0494)	0.0747(0.0079)	0.0737(0.0070)	0.0747(0.0079)	0.0738(0.0070)
	150	20	150	0.8342(0.0503)	0.0606(0.0059)	0.0599(0.0053)	0.0606(0.0059)	0.0599(0.0053)
	200	20	200	0.8394(0.0499)	0.0521(0.0049)	0.0515(0.0044)	0.0521(0.0049)	0.0515(0.0044)
				设置 B： $p_2 = 20$、 $p_1 = T$				
$\mathcal{D}$(R, R)	20	20	20	0.6178(0.1205)	0.0954(0.0173)	0.0938(0.0158)	0.1151(0.0308)	0.0947(0.0162)
	50	50	20	0.5514(0.1440)	0.0574(0.0058)	0.0569(0.0056)	0.0595(0.0061)	0.0574(0.0058)
	100	100	20	0.4945(0.1526)	0.0402(0.0037)	0.0399(0.0036)	0.0406(0.0038)	0.0402(0.0037)
	150	150	20	0.4481(0.1573)	0.0327(0.0027)	0.0325(0.0027)	0.0328(0.0027)	0.0327(0.0027)
	200	200	20	0.4193(0.1671)	0.0281(0.0021)	0.0279(0.0021)	0.0281(0.0021)	0.0281(0.0021)
$\mathcal{D}$(C, C)	20	20	20	0.6138(0.1182)	0.0947(0.0166)	0.0933(0.0156)	0.1148(0.0290)	0.0942(0.0160)
	50	50	20	0.5903(0.1269)	0.0359(0.0054)	0.0357(0.0053)	0.0599(0.0194)	0.0357(0.0053)
	100	100	20	0.5797(0.1291)	0.0177(0.0025)	0.0177(0.0025)	0.0469(0.0206)	0.0177(0.0025)
	150	150	20	0.5728(0.1292)	0.0116(0.0017)	0.0116(0.0016)	0.0424(0.0185)	0.0116(0.0016)
	200	200	20	0.5761(0.1287)	0.0088(0.0012)	0.0088(0.0012)	0.0425(0.0204)	0.0088(0.0012)
$\mathcal{D}$(Vec(F_t), Vec(F_t))	20	20	20	0.6740(0.0380)	0.1837(0.0371)	0.1783(0.0338)	0.1837(0.0371)	0.1784(0.0339)
	50	50	20	0.7925(0.0454)	0.1080(0.0123)	0.1064(0.0113)	0.1080(0.0123)	0.1064(0.0113)
	100	100	20	0.8253(0.0480)	0.0751(0.0073)	0.0742(0.0066)	0.0751(0.0073)	0.0742(0.0066)
	150	150	20	0.8349(0.0466)	0.0604(0.0061)	0.0597(0.0057)	0.0604(0.0061)	0.0597(0.0057)
	200	200	20	0.8382(0.0504)	0.0524(0.0051)	0.0518(0.0047)	0.0524(0.0051)	0.0518(0.0047)

Lam （2021）提出的 TCorTh 方法在估计因子数量对方面的表现。表2 展示了不同方法在情景A和情景B下进行500次重复试验时，精确估计和低估的频率。我们采用与第4.1 节介绍的相同数据生成方案，但设定 $k_1 = 3$， $k_2 = 2$。并对所有方法设定 $k_{\max} = 8$ 。从表2可以看出，RPILS-ER和IterER的表现相当，且都优于其他方法。随着 p_1 和 p_2 增大，除TCorTh方法外，所有方法的表现均有所提升。还可以看出，随着 p_1 和 p_2 的维度增大，RPILS-ER精确估计的比例趋于1，这与我们的理论分析一致。

表2：设置A和设置B下，500次重复试验中因子数量的精确估计和低估频率。

T	RPILS-ER	IterER	α-主成分分析-ER	iTOP-ER	iTIP-ER	TCorTh
情景A： $p_1 = 20, p_2 = T, k_1 = 3, k_2 = 2, k_{\max} = 8$						
20	0.6880(0.3120)	0.9800(0.0200)	0.4660(0.5340)	0.5920(0.4080)	0.1160(0.8840)	0.5100(0.4900)
50	0.9840(0.0160)	1.0000(0.0000)	0.7040(0.2960)	0.8600(0.1400)	0.1880(0.8120)	0.9560(0.0440)
100	0.9940(0.0060)	1.0000(0.0000)	0.6900(0.3100)	0.9240(0.0760)	0.4140(0.5860)	0.9900(0.0100)
150	1.0000(0.0000)	1.0000(0.0000)	0.7040(0.2960)	0.9680(0.0320)	0.5340(0.4660)	0.9880(0.0120)
200	1.0000(0.0000)	1.0000(0.0000)	0.6960(0.3040)	0.9800(0.0200)	0.7200(0.2800)	0.9400(0.0600)
情景B: $p_2 = 20, p_1 = T, k_1 = 3, k_2 = 2, k_{\max} = 8$						
20	0.6880(0.3120)	0.9800(0.0200)	0.4660(0.5340)	0.5920(0.4080)	0.1160(0.8840)	0.5100(0.4900)
50	0.9980(0.0020)	1.0000(0.0000)	0.9440(0.0560)	0.8980(0.1020)	0.2360(0.7640)	1.0000(0.0000)
100	1.0000(0.0000)	1.0000(0.0000)	0.9600(0.0400)	0.9720(0.0280)	0.3560(0.6440)	0.9760(0.0240)
150	1.0000(0.0000)	1.0000(0.0000)	0.9640(0.0360)	0.9860(0.0140)	0.4760(0.5240)	0.8600(0.1400)
200	1.0000(0.0000)	1.0000(0.0000)	0.9400(0.0600)	0.9940(0.0060)	0.6100(0.3900)	0.7200(0.2800)

5 真实数据示例

5.1 Fama-French 10×10 投资组合

本节研究 Wang 等人 (2019) 和 Yu 等人 (2022) 研究的一个金融投资组合数据集。该数据集由 100 个投资组合的月收益率构成，每个时点均以结构规整的 10×10 矩阵表示，行对应市场资本规模的 10 个水平（记为 S1-S10），列对应账面市值比的 10 个水平（记为 BE1-BE10）。该数据集收录了从 1964 年 1 月至 2019 年 12 月的月收益率，共计 672个月。详情请访问网站 http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html。按照

与 Yu 等人 (2022) 和 Wang 等人 (2019) 相同的预处理方法，我们先减去相应的月度市场超额收益，再对每条收益序列进行标准化，从而调整收益序列。我们采用 Xiong 和 Pelger (2022) 提出的基于因子模型的方法填补缺失值。所有增广迪基-富勒检验均拒绝原假设，这表明所有序列均具有平稳性。

关于因子数，RPILS-ER、 α-PCA-ER、iTOP-IC 和 iTIP-IC 均建议 $(k_1, k_2) = (1, 1)$，而 IterER 建议 $(k_1, k_2) = (2, 1)$；iTOP-ER、iTIP-ER 和 TCorTh 均建议 $(k_1, k_2) = (2, 2)$。为便于说明，我们采用 Yu 等（2022）中的 $(k_1, k_2) = (2, 2)$。方差最大旋转和缩放后的估计载荷矩阵列于表 3。

表 3 展示了方差最大旋转和缩放后的估计行载荷矩阵与列载荷矩阵。从表中可以看出，所提出的 RPILS 方法在估计载荷矩阵方面与 PE、 α-PCA 和 ACCE 方法表现相近。从规模角度看，小规模投资组合在第一个因子上的载荷较大，而大规模投资组合在第二个因子上的载荷较大。从账面市值比角度看，低账面市值比投资组合在第二个因子上的载荷较大，而高账面市值比投资组合的载荷主要集中在第一个因子上。

表 3：Fama-French 数据集经方差最大旋转并缩放 30 倍后的载荷矩阵。

		Size									
方法	因子	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
RPILS	1	16	15	12	10	8	5	3	1	-4	-7
	2	5	1	-3	-5	-8	-10	-12	-13	-15	-11
PE	1	-16	-15	-12	-10	-8	-5	-3	-1	4	7
	2	-6	-1	3	5	8	10	12	13	15	11
α-主成分分析	1	-14	-14	-13	-11	-9	-7	-4	-2	3	7
	2	-4	-2	1	3	6	9	12	13	16	14
ACCE	1	-12	-14	-12	-13	-10	-6	-3	-1	4	9
	2	-1	-1	-1	2	5	10	11	18	15	11
		账面市值比									
方法因子		BE1	BE2	BE3	BE4	BE5	BE6	BE7	BE8	BE9	BE10
RPILS	1	-6	-1	4	7	10	11	12	12	12	10
	2	20	17	11	8	4	2	0	-1	-1	0
PE	1	6	1	-4	-7	-10	-11	-12	-12	-12	-10
	2	20	17	11	8	4	2	0	-1	-1	0
α-主成分分析	1	6	2	-4	-7	-10	-11	-12	-13	-12	-11
	2	19	18	12	8	4	2	0	-1	-1	-1
ACCE	1	6	-1	-4	-8	-8	-9	-10	-13	-15	-12
	2	21	15	11	6	5	2	1	-2	-3	1

我们还采用与Yu 等人（2022）和Wang 等人（2019）相同的滚动验证方案，进一步比较这些方法。对于每一年 t 从 1996 年到 2019 年，我们反复使用 n 年前（带宽） t 的数据来拟合

表4：法玛-弗伦奇投资组合的滚动验证。训练集的样本量为 $12n$ 和 $k_1 = k_2 = k$。$\overline{\text{MSE}}$ 均方误差 $\bar{\rho}$，$\bar{v}$ 分别为平均定价误差、总方差未解释比例的平均值和估计载荷空间的平均变化。

		$\overline{\text{MSE}}$				$\bar{\rho}$				$\bar{v}$			
n	k	RPILS	PE	α-主成分分析	ACCE	RPILS	PE	α-主成分分析	ACCE	RPILS	PE	α-主成分分析	ACCE
5	1	0.8766	0.8703	**0.8624**	0.8846	0.8001	0.8022	**0.7960**	0.8284	0.2653	**0.1757**	0.2412	0.3032
10	1	0.8735	**0.8548**	0.8596	0.8797	0.7948	**0.7836**	0.7913	0.8149	0.2341	**0.0847**	0.2027	0.1654
15	1	0.8570	**0.8530**	0.8599	0.8844	0.7828	**0.7822**	0.7918	0.8118	**0.0606**	0.0636	0.2335	0.1520
5	2	**0.5954**	0.5965	0.6010	0.6673	**0.6231**	0.6248	0.6284	0.6727	**0.2331**	0.2390	0.3497	0.4605
10	2	0.6014	**0.6013**	0.6108	0.6545	**0.6273**	0.6276	0.6364	0.6684	**0.0892**	0.0924	0.2606	0.2568
15	2	0.6027	**0.6025**	0.6115	0.6375	**0.6261**	0.6262	0.6302	0.6516	**0.0564**	0.0573	0.1735	0.1894
5	3	**0.5204**	0.5216	0.5291	0.5639	**0.5473**	0.5495	0.5558	0.5900	**0.2781**	0.2865	0.4321	0.4974
10	3	**0.5181**	0.5193	0.5262	0.5728	**0.5465**	0.5481	0.5549	0.5936	**0.1074**	0.1142	0.3532	0.3029
15	3	**0.5166**	0.5172	0.5220	0.5601	**0.5438**	0.5446	0.5444	0.5825	**0.0783**	0.0839	0.3082	0.2986

矩阵变量因子模型，并估计载荷矩阵。随后使用估计得到的载荷来估计当前年份 12 个月的因子及相应残差。具体而言，设 $\mathbf{Y}_t^i$ 和 $\widehat{\mathbf{Y}}_t^i$ 分别为年份 t中月份 i 的观测价格矩阵和估计价格矩阵，将 $\bar{\mathbf{Y}}_t$ 记为平均价格矩阵，并定义

$$\text{MSE}_t = \frac{1}{12 \times 10 \times 10} \sum_{i=1}^{12} \|\widehat{\mathbf{Y}}_t^i - \mathbf{Y}_t^i\|_F^2, \quad \rho_t = \frac{\sum_{i=1}^{12} \|\widehat{\mathbf{Y}}_t^i - \mathbf{Y}_t^i\|_F^2}{\sum_{i=1}^{12} \|\mathbf{Y}_t^i - \bar{\mathbf{Y}}_t\|_F^2},$$

分别作为平均平方定价误差和总方差的未解释比例。载荷空间的变化通过

$$v_t = \mathcal{D}(\widehat{\mathbf{C}}_t \otimes \widehat{\mathbf{R}}_t, \widehat{\mathbf{C}}_{t-1} \otimes \widehat{\mathbf{R}}_{t-1}),$$

在滚动验证程序中。

表 4 报告了所提出的 RPILS 方法和其他方法的均方误差均值结果 ρ，v针对不同的带宽 n 和因子数量 k_1，k_2就平均均方误差而言，我们的 RPILS 方法与 PE 方法相当，且优于其他方法 ρ，v。

5.2 跨国宏观经济指数

本节分析一个从经济合作与发展组织（经合组织）收集的跨国宏观经济指数数据集。该数据集包含8个国家在1988年第一季度至2020年第二季度共130个季度的10项宏观经济指数。这8个国家分别是美国、英国、加拿大、法国、德国、挪威、澳大利亚和新西兰。这些指数分属4个主要类别，即消费价格、利率、生产和国际贸易。有关该数据集的预处理过程，详见Yu 等人（2022）。

关于因子数，RPILS-ER和 α-主成分分析建议（k_1，k_2）=（1，2），而IterER建议（k_1，k_2）=（1，5）。为便于说明，我们采纳k_1，k_2）=（3，4）的建议，该建议出自Yu 等人（2022）。估计得到的载荷矩阵见表5 和表6。所提出的RPILS方法与PE方法的表现几乎相同。正如Yu 等人（2022）总结的那样，除德国以外的国家自然分为3组：大洋洲、北美洲和欧洲。另一方面，宏观经济指数分为4组：消费价格、利率、生产和国际贸易，这与经济解释相吻合。

表5：跨国宏观经济指数数据集中按不同方法得到的行载荷矩阵，经方差最大旋转并乘以10。

方法	因子	澳大利亚	新西兰	美国	加拿大	挪威	德国	法国	英国
RPILS	1	5	-4	1	-4	-20	-7	-15	-12
	2	0	1	-20	-16	11	-8	-5	-4
	3	24	16	1	-3	0	6	-1	2
PE	1	0	1	-7	-6	3	-3	-2	-1
	2	2	-2	1	-1	-7	-2	-5	-5
	3	8	6	0	-1	0	2	-1	1
α-主成分分析	1	-1	1	-7	-5	3	-3	-2	-1
	2	1	-1	0	-1	-7	-2	-5	-4
	3	-7	-7	0	1	0	-1	1	-1
ACCE	1	2	-2	1	-2	-6	0	-6	-5
	2	7	5	0	0	0	5	0	0
	3	0	-2	8	4	-2	1	1	2

表6：跨国宏观经济指数数据集中按不同方法得到的列载荷矩阵，经方差最大旋转并乘以10。

方法	因子	CPI:Tot	CPI：能源	CPI：非食品非能源	IR:3-Mon	IR:Long	P:TIEC	P:TM	GDP	IT:Ex	IT:Im
RPILS	1	-2	4	-7	-3	3	-20	-20	-5	2	0
	2	1	1	-2	-18	-24	1	-1	2	1	-3
	3	-2	6	-7	1	0	1	1	16	19	14
	4	-19	-20	-10	4	-4	1	-1	-1	-1	1
PE	1	1	-2	3	1	-1	6	7	2	-1	0
	2	6	7	3	-1	1	0	0	0	0	0
	3	0	0	-1	-6	-8	0	0	1	0	-1
	4	1	-2	3	0	0	-1	0	-5	-6	-5
α-主成分分析	1	0	-1	1	1	-1	7	6	4	0	0
	2	7	5	5	-1	1	0	1	0	0	0
	3	0	0	0	-7	-7	1	0	-1	1	0
	4	0	2	-2	0	0	0	0	2	7	6
ACCE	1	0	0	0	0	1	-7	-7	0	0	0
	2	1	0	0	-5	-4	1	-1	-2	-4	-6
	3	-4	2	-9	0	0	2	-1	2	0	0
	4	6	7	0	-1	3	0	0	2	1	-1

我们还采用滚动预测程序，以进一步考察不同方法的实际应用。首先，考虑所选国家在时点 t 的通货膨胀变化（居民消费价格总指数的对数水平的二阶差分，CPI:Tot），记为 y_t。令 $\boldsymbol{x}_t$ 为所选国家在时点 t 的其余 9 个指数构成的向量，并令 $\mathbf{Z}_t$ 为时点 8×10 的面板数据 t，其行对应各个国家，列对应所有宏观经济指数。我们采用以下自回归（AR）模型（模型1）和因子增强自回归（FAAR）模型（模型2–4）预测 y_{t+1}，这与 Stock 和 Watson (2002b) 的扩散指数预测类似。

模型 **1** $y_{t+1} = a + by_t + \epsilon_{t+1}$,

模型 **2** $y_{t+1} = a + by_t + \boldsymbol{\beta}^\top \boldsymbol{f}_{1t} + \epsilon_{t+1}$，其中 $\boldsymbol{f}_{1t}$ 是根据观测值为 $\{\boldsymbol{x}_t\}$ 的向量因子模型估计得到的。

模型 **3** $y_{t+1} = a + by_t + \boldsymbol{\beta}^\top \boldsymbol{f}_{2t} + \epsilon_{t+1}$，其中 $\boldsymbol{f}_{2t}$ 是基于观测值 $\{\text{Vec}(\mathbf{Z}_t)\}$，通过向量因子模型估计得到的。

模型 **4** $y_{t+1} = a + by_t + \boldsymbol{\beta}^\top \text{Vec}(\mathbf{F}_t) + \epsilon_{t+1}$，其中 $\mathbf{F}_t$ 是分别通过 RPILS、PE、ACCE 和 α-主成分分析，基于观测值 $\{\mathbf{Z}_t\}$，利用矩阵因子模型估计得到的。

这里用于比较的模型与Yu 等人(2022)中的模型完全相同。为完整起见，我们在此再次说明这些模型。首先，模型1是一个简单的自回归模型。模型2在模型1的自回归模型中加入所选国家的共同指数因子。在模型3和模型4中，指数因子和国家因子都纳入考量。两者的区别在于，模型4采用更简约的矩阵因子结构，而模型3则将矩阵时间序列向量化，并采用向量因子结构。为避免预测中可能出现的过拟合，我们还使用 LASSO（Tibshirani，1996）选择因子，并估计模型2至模型4的系数。

对于每个季度 t，我们使用 t 之前的 80 个相邻观测值训练模型，并预测 y_{t+1} (记作 $\widehat{y}_{t+1}$)。由于 y_t 在预处理中已标准化，我们通过乘以标准差并加回样本均值，将预测的 y_{t+1} 转换为与通货膨胀率的变化相匹配。为简化记号，我们仍将转换后的预测变量记作 $\widehat{y}_{t+1}$。$\widehat{}$通货膨胀率 I_{t+1} 随后通过整合 $\widehat{y}_{t+1}$ 和 I_t进行预测，即 $I_{t+1} = \widehat{y}_{t+1} + I_t$。在模型2和模型3中，变量选择前的因子数分别设为 k_2 和 $k_1 \times k_2$。我们还重点考察（k_1，k_2）=（3，4）的情形。表7 上半部分展示了年化通胀率的平均绝对百分比误差（MAPE）。对于大洋洲最大的国家澳大利亚，采用 RPILS 的模型4在 MAPE 指标上预测表现最佳。对于挪威和新西兰，模型1表现最佳，这表明指数因子和国家因子在模型2至模型4中起到了噪声的作用。对于美国、加拿大、法国、大不列颠和德国，指数因子和国家因子都有助于提升预测表现。

表7：不同国家采用不同方法得到的通货膨胀率和GDP增长率（均为年率）的平均绝对百分比误差，$(k_1, k_2) = (3, 4)$。

模型	AUS	NZL	USA	CAN	NOR	DEU	FRA	GBR
通货膨胀率的平均绝对百分比误差								
模型1	1.5880	**1.8866**	2.8346	2.4605	**1.9359**	1.8751	1.8060	1.5776
模型2	1.6258	1.9086	2.5835	2.2349	2.2607	1.9574	1.8322	1.5019
模型3	1.5989	2.0507	2.4709	1.9653	2.3865	**1.7279**	**1.2411**	**1.1980**
模型4（RPILS）	**1.5607**	1.9319	2.6829	2.2736	2.1377	1.8756	1.7236	1.3863
模型4（PE）	1.5886	2.0038	2.3527	1.8515	2.1550	**1.7330**	1.4074	1.3241
模型4（ACCE）	1.5853	1.9292	**2.2466**	**1.8273**	2.4442	1.8040	1.3822	1.3348
模型4（α-主成分分析）	1.5981	1.9177	2.3777	1.9023	2.3132	1.7402	1.4335	1.3490
GDP增长率的平均绝对百分比误差								
模型1	1.8897	2.9211	**2.4317**	2.8279	3.7794	**3.2970**	**2.6777**	**3.2515**
模型2	**1.8709**	2.9117	2.4683	2.8584	3.8977	3.3413	2.8395	3.3231
模型3	**1.8669**	**2.8391**	2.5664	2.6364	3.7375	3.9167	2.8535	3.3285
模型4（RPILS）	1.8904	3.0531	2.5087	2.7084	3.7467	**3.3315**	2.7672	3.3198
模型4（市盈率）	1.8731	2.9268	2.5819	**2.5505**	3.9012	3.4485	2.8079	3.3523
模型4（ACCE）	1.9118	3.0370	2.5067	2.6869	**3.6739**	3.5838	2.7446	3.4251
模型4（α-主成分分析）	**1.8709**	3.0364	2.6078	2.6544	3.8955	3.5226	2.8701	3.3843

请注意，结果还表明，矩阵因子结构可进一步改善对两个美洲国家的预测。我们还对所有国家的国内生产总值增长率（国内生产总值对数水平的一阶差分）进行滚动预测。表7 下半部分所示结果表明，对于制造业实力强劲的美洲和欧洲国家——美国、德国、法国和大不列颠，简单的自回归模型足以准确预测国内生产总值增长率。对于其他国家，国家因子和指数因子有助于提升国内生产总值增长率的预测表现；而对于加拿大和挪威，矩阵因子结构的优势更为明显。

6 讨论

我们提出了一种用于矩阵因子模型的简易迭代最小二乘算法。第一步，通过使用两个确定性权重矩阵对观测值进行投影，估计潜在因子矩阵。我们证明，即使

行因子数和列因子数都被高估，在满足某些正则性条件时，对因子的推断仍然渐近有效。第二步，在满足某些可识别性条件的情况下，通过最小化平方弗罗贝尼乌斯损失函数来估计行/列载荷矩阵。随后，将得到的载荷矩阵估计量用作新的权重/投影矩阵，并反复执行上述两个步骤直至收敛。给定因子矩阵的真实维度，我们建立了任意 $s \geq 1$ 下第 s 次迭代所得载荷矩阵和公共成分估计量的收敛速率。为确定因子数对，我们提出了一种基于迭代结果的特征值比率法，并证明了由此得到的估计量具有一致性。

作为未来的一个研究方向，我们的方法可以推广到张量值时间序列。直观地说，如果将（3.1）中的平方损失替换为Huber损失，就能得到更稳健的估计量。这一点本身也很值得研究，因为现实中的金融收益和宏观经济指数往往呈现厚尾特征。由于还需要开展大量额外工作，我们将此留待未来研究。

参考文献

Ahn, S.C., Horenstein, A.R., 2013年。用于确定因子数量的特征值比率检验。《Econometrica》81, 1203–1227。

Aït-Sahalia, Y., Xiu, D., 2017。使用主成分分析估计包含高频数据的高维因子模型。《计量经济学杂志》201，388–399。

Bai, J., 2003年。大维度因子模型的推断理论。《Econometrica》71，135–171。

Barigozzi, M., Cho, H., 2020年. 因子数量被高估时高维因子模型的一致估计. Electronic Journal of Statistics 14, 2892–2921.

Barigozzi, M., Cho, H., Fryzlewicz, P., 2018. 高维时间序列的多重变点与因子分析. 《计量经济学杂志》206, 187–225.

Chen, E.Y., Fan, J., 2021. 高维矩阵变量因子模型的统计推断. Journal of the American Statistical Association（刚刚接收）, 1–44. Chen, L., Dolado, J.J., Gonzalo, J., 2021. 分位数因子模型. Econometrica 89, 875–910.

Fan, J., Liao, Y., 2022. 从多样化投影中学习潜在因子及其在过度估计因子和弱因子中的应用. Journal of the American Statistical Association 117, 909–924.

Fan, J., Liao, Y., Mincheva, M., 2013年. 通过阈值化主正交补进行大规模协方差估计. 《皇家统计学会杂志：B辑（统计方法学）》75, 603–680.

Han, Y., Chen, R., Zhang, C.H., 2022. 张量因子模型中的秩确定. Electronic Journal of Statistics 16, 1726–1803. He, Y., Kong, X., , Yu, L., Zhang, X., Zhao, C., 2021a. 矩阵因子分析：从最小二乘到迭代投影. arXiv: 2112.04186 . He, Y., Kong, X., Trapani, L., Yu, L., 2021b. 矩阵序列的单向或双向因子模型? arXiv: arXiv:2110.01008 . He, Y., Kong, X., Yu, L., Zhang, X., 2022a. 无矩约束的大维因子分析. Journal of Business & Economic Statistics 40, 302–312. He, Y., Wang, Y., , Yu, L., Zhou, W., Zhou, W., 2022b. 高维矩阵 Kendall's tau：矩阵因子模型的稳健统计量. arXiv e-prints: 2207.09633 . Kong, X., 2017. 基于高频数据确定公共因子的数量. Biometrika 104, 397–410. Lam, C., 2021. 使用相关性阈值法确定时间序列张量因子模型的秩. 技术报告. LSE 工作论文. Lam, C., Yao, Q., 2012. 高维时间序列的因子建模：因子数量的推断. The Annals of Statistics 40, 694–726. Moon, H.R., Weidner, M., 2015. 将因子数量未知的面板数据线性回归视为交互固定效应. Econometrica 83, 1543–1579. Onatski, A., 2009. 大型因子模型中关于因子数量的假设检验. Econometrica 77, 1447–1479. Stock, J.H., Watson, M.W., 2002a. 利用大量预测变量的主成分进行预测. Journal of the American Statistical Association 97, 1167–1179. Stock, J.H., Watson, M.W., 2002b. 使用扩散指数进行宏观经济预测. Journal of Business & Economic Statistics 20, 147–162. Tibshirani, R., 1996. 通过 LASSO 进行回归收缩与变量选择. 《皇家统计学会杂志：B辑（方法学）》58, 267–288. Trapani, L., 2018. 用于确定因子数量的随机序贯程序. Journal of the American Statistical Association 113, 1341–1349. Wang, D., Liu, X., Chen, R., 2019. 矩阵值高维时间序列的因子模型. 《计量经济学杂志》208, 231–248.

Westerlund, J., Urbain, J.P., 2015. 截面均值与主成分。《计量经济学杂志》185, 372–377. Xiong, R., Pelger, M., 2022. 基于缺失观测的高维潜因子建模及其在因果推断中的应用。《计量经济学杂志》，即将刊出。Yu, L., He, Y., Kong, X., Zhang, X., 2022. 高维矩阵因子模型的投影估计。《计量经济学杂志》229, 201–217. Yu, L., He, Y., Zhang, X., 2019. 高维椭圆因子模型的稳健因子数设定。《多元分析杂志》174, 104543.

《通过随机投影实现大维矩阵因子模型的高效迭代最小二乘算法》补充材料

Yong He *、Ran Zhao*、Wen-Xin Zhou†、

本文提　　供了主要定理的详细证明，以及额外的引理和命题　　.

S1 主要定理的证明

S1.1 定理 2.1 的证明

证明。由 $\mathrm{H}_1^\top(\mathrm{H}_1\mathrm{H}_1^\top)^+\mathrm{H}_1 = \mathrm{I}_{k_1}$、$\mathrm{H}_2^\top(\mathrm{H}_2\mathrm{H}_2^\top)^+\mathrm{H}_2 = \mathrm{I}_{k_2}$，可得 $\widehat{\mathrm{F}}_t = \mathrm{H}_1\mathrm{F}_t\mathrm{H}_2^\top + \mathrm{E}_t$，这意味着 $\mathrm{M}_1^\top\widehat{\mathrm{F}}_t\mathrm{M}_2 - \mathrm{F}_t = \mathrm{M}_1^\top\mathrm{E}_t\mathrm{M}_2$，其中 $\mathrm{M}_1 = (\mathrm{H}_1\mathrm{H}_1^\top)^+\mathrm{H}_1$,$\mathrm{M}_2 = (\mathrm{H}_2\mathrm{H}_2^\top)^+\mathrm{H}_2$。由于 $\|(\mathrm{H}_1\mathrm{H}_1^\top)^+\mathrm{H}_1\|_2 = O_p(\nu^{-1}\min(\mathrm{H}_1))$ 且 $\|(\mathrm{H}_2\mathrm{H}_2^\top)^+\mathrm{H}_2\|_2 = O_p(\nu^{-1}\min(\mathrm{H}_2))$，由引理 S2.2 (1)，可得

$$\|\mathbf{M}_1^\top\widehat{\mathbf{F}}_t\mathbf{M}_2 - \mathbf{F}_t\|_2 = O_p\left(\frac{1}{\sqrt{p_1}}\nu_{\min}^{-1}(\mathbf{H}_1)\nu_{\min}^{-1}(\mathbf{H}_2)\right).$$

□

S1.2 定理 2.2的证明

证明。根据命题 S2.1，$\lambda_{\min}(\frac{1}{T}\mathrm{M}^\top\widehat{\mathrm{F}}^\top\widehat{\mathrm{F}}\mathrm{M}) \geq \lambda_{\min}(\frac{1}{T}\widehat{\mathrm{F}}^\top\widehat{\mathrm{F}})\lambda_{\min}(\mathrm{M}^\top\mathrm{M}) \geq c(p_1p_2)^{-1}\lambda_{\min}(\boldsymbol{D}^{-2}\mathbf{H})$ 以高概率成立。根据 $\mathrm{H}^\top$ 的奇异值分解，即 $\mathrm{H}^\top = \mathrm{U}_{\mathbf{H}}(\mathrm{D}_{\mathbf{H}}, 0)\mathrm{E}_{\mathbf{H}}^\top$，我们得出 $\mathrm{P}_{\widehat{\mathbf{F}}\mathbf{M}}$ 有明确定义。

由于 $\widehat{\mathrm{F}} = \mathrm{F}\mathrm{H}^\top + \mathrm{E}$，因此有 $\widehat{\mathrm{F}}\mathrm{M} - \mathrm{F} = \mathrm{E}(\mathrm{H}\mathrm{H}^\top)^+\mathrm{H}$，其中 $\mathrm{M} = (\mathrm{H}\mathrm{H}^\top)^+\mathrm{H}$。进一步根据 $\|(\mathrm{H}\mathrm{H}^\top)^+\mathrm{H}\|_2 = O_p\left(\nu^{-1}\min\right)$ 这一事实以及引理 S2.2 (2)、(3)，我们有

$$\frac{1}{\sqrt{T}}\|\widehat{\mathbf{F}}\mathbf{M} - \mathbf{F}\|_2 = \frac{1}{\sqrt{T}}\|\boldsymbol{\mathcal{E}}(\mathbf{H}\mathbf{H}^\top)^+\mathbf{H}\|_2 \leq \frac{1}{\sqrt{T}}\|\boldsymbol{\mathcal{E}}\|_2\|(\mathbf{H}\mathbf{H}^\top)^+\mathbf{H}\|_2 = O_p\left(\frac{1}{\sqrt{p_1p_2}}\nu_{\min}^{-1}\right),$$

$$\frac{1}{T}\|\mathbf{F}^\top(\widehat{\mathbf{F}}\mathbf{M} - \mathbf{F})\|_2 \leq \frac{1}{T}\|\mathbf{F}^\top\boldsymbol{\mathcal{E}}\|_2\|(\mathbf{H}\mathbf{H}^\top)^+\mathbf{H}\|_2 = O_p\left(\frac{1}{\sqrt{Tp_1p_2}}\nu_{\min}^{-1}\right).$$

进一步根据

$$\begin{aligned}\left\|\frac{1}{T}\mathbf{M}^\top\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}\mathbf{M} - \frac{1}{T}\mathbf{F}^\top\mathbf{F}\right\|_2 &= \left\|\frac{1}{T}(\widehat{\mathbf{F}}\mathbf{M} - \mathbf{F})^\top(\widehat{\mathbf{F}}\mathbf{M} - \mathbf{F}) + \frac{1}{T}\mathbf{F}^\top(\widehat{\mathbf{F}}\mathbf{M} - \mathbf{F}) + \frac{1}{T}(\widehat{\mathbf{F}}\mathbf{M} - \mathbf{F})^\top\mathbf{F}\right\|_2 \\ &\leq \frac{1}{T}\|\widehat{\mathbf{F}}\mathbf{M} - \mathbf{F}\|_2^2 + \frac{2}{T}\|\mathbf{F}^\top(\widehat{\mathbf{F}}\mathbf{M} - \mathbf{F})\|_2 = O_p\left(\frac{1}{\sqrt{Tp_1p_2}}\nu_{\min}^{-1} + \frac{1}{p_1p_2}\nu_{\min}^{-2}\right),\end{aligned}$$

我们有

$$\left\|(\frac{1}{T}\mathbf{M}^\top\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}\mathbf{M})^{-1}\right\|_2 = O_p(1)$$

*山东大学金融研究院，中国。电子邮箱：heyong@sdu.edu.cn，　Zhaoran@mail.sdu.edu.cn †美国加利福尼亚大学圣迭戈分校数学科学系。电子邮箱：wez243@ucsd.edu

并且由于 $\mathsf{A}^{-1}-\mathsf{B}^{-1}=\mathsf{A}^{-1}(\mathsf{B}-\mathsf{A})\mathsf{B}^{-1}$，我们得到

$$\left\|\left(\frac{1}{T}\mathbf{M}^\top\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}\mathbf{M}\right)^{-1}-\left(\frac{1}{T}\mathbf{F}^\top\mathbf{F}\right)^{-1}\right\|_2=O_p\left(\frac{1}{\sqrt{Tp_1p_2}}\nu_{\min}^{-1}+\frac{1}{p_1p_2}\nu_{\min}^{-2}\right).$$

因此，

$$\begin{aligned}\mathbf{P}_{\widehat{\mathbf{F}}\mathbf{M}}-\mathbf{P}_{\mathbf{F}}=&\frac{1}{\sqrt{T}}\widehat{\mathbf{F}}\mathbf{M}\left[(\frac{1}{T}\mathbf{M}^\top\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}\mathbf{M})^{-1}-(\frac{1}{T}\mathbf{F}^\top\mathbf{F})^{-1}\right](\frac{1}{\sqrt{T}}\widehat{\mathbf{F}}\mathbf{M})^\top\\&+\frac{1}{\sqrt{T}}\widehat{\mathbf{F}}\mathbf{M}(\frac{1}{T}\mathbf{F}^\top\mathbf{F})^{-1}\frac{1}{\sqrt{T}}(\widehat{\mathbf{F}}\mathbf{M}-\mathbf{F})^\top+\frac{1}{\sqrt{T}}(\widehat{\mathbf{F}}\mathbf{M}-\mathbf{F})(\frac{1}{T}\mathbf{F}^\top\mathbf{F})^{-1}\frac{1}{\sqrt{T}}\mathbf{F}^\top,\end{aligned}$$

$$\|\mathbf{P}_{\widehat{\mathbf{F}}\mathbf{M}}-\mathbf{P}_{\mathbf{F}}\|_2=O_p\left(\frac{1}{\sqrt{p_1p_2}}\nu_{\min}^{-1}\right).$$

最后，注意到 $\mathsf{P}_{\widehat{\mathbf{F}}}\mathsf{P}_{\widehat{\mathbf{F}}\mathbf{M}}=\mathsf{P}_{\widehat{\mathbf{F}}\mathbf{M}}$，我们得到

$$\|\mathbf{P}_{\widehat{\mathbf{F}}}\mathbf{P}_{\mathbf{F}}-\mathbf{P}_{\mathbf{F}}\|_2\le\|\mathbf{P}_{\widehat{\mathbf{F}}}(\mathbf{P}_{\mathbf{F}}-\mathbf{P}_{\widehat{\mathbf{F}}\mathbf{M}})\|_2+\|\mathbf{P}_{\widehat{\mathbf{F}}\mathbf{M}}-\mathbf{P}_{\mathbf{F}}\|_2=O_p\left(\frac{1}{\sqrt{p_1p_2}}\nu_{\min}^{-1}\right).$$

□

S1.3 定理的证明 3.1

证明。以下不失一般性地假设 $m_1=k_1=1,m_2=k_2=1$.

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^T\mathbf{X}_t\mathbf{W}_2\widehat{\mathbf{F}}_t^\top&=\frac{1}{Tp_1p_2}\sum_{t=1}^T\mathbf{X}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{X}_t^\top\mathbf{W}_1\\&=\frac{1}{Tp_1p_2}\sum_{t=1}^T(\mathbf{R}\mathbf{F}_t\mathbf{C}^\top+\mathbf{E}_t)\mathbf{W}_2\mathbf{W}_2^\top(\mathbf{R}\mathbf{F}_t\mathbf{C}^\top+\mathbf{E}_t)^\top\mathbf{W}_1\\&=\frac{1}{Tp_1p_2}\sum_{t=1}^T\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1+\frac{1}{Tp_1p_2}\sum_{t=1}^T\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\\&+\frac{1}{Tp_1p_2}\sum_{t=1}^T\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1+\frac{1}{Tp_1p_2}\sum_{t=1}^T\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\\&:=\boldsymbol{\delta}_1+\boldsymbol{\delta}_2+\boldsymbol{\delta}_3+\boldsymbol{\delta}_4.\end{aligned}$$

下面我们逐项分析 $\boldsymbol{\delta}_1,\boldsymbol{\delta}_2,\boldsymbol{\delta}_3,\boldsymbol{\delta}_4$。

对于第一项 $\boldsymbol{\delta}_1$，一方面，我们有

$$\begin{aligned}\|\boldsymbol{\delta}_1\|_F&=\left\|\frac{1}{Tp_1p_2}\sum_{t=1}^T\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\right\|_F=\left\|\frac{p_2}{T}\sum_{t=1}^T\mathbf{R}\mathbf{F}_t\mathbf{H}_2^\top\mathbf{H}_2\mathbf{F}_t^\top\mathbf{H}_1\right\|_F\\&\le p_2\|\mathbf{R}\|_F\left\|\frac{1}{T}\sum_{t=1}^T\mathbf{F}_t\mathbf{F}_t^\top\right\|_F\|\|\mathbf{H}_2\|_2^2\|\mathbf{H}_1\|_2=O_p\left(\sqrt{p_1}p_2\nu_{\max}(\mathbf{H}_1)\nu_{\max}^2(\mathbf{H}_2)\right)\\&=O_p\left(\sqrt{p_1}p_2\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)\right).\end{aligned}$$

另一方面，

$$\|\boldsymbol{\delta}_1\|_F \geq \|\boldsymbol{\delta}_1\|_2 = \left\| \frac{p_2}{T} \sum_{t=1}^{T} \mathbf{R}\mathbf{F}_t \mathbf{H}_2^\top \mathbf{H}_2 \mathbf{F}_t^\top \mathbf{H}_1 \right\|_2 = \sqrt{p_1} p_2 \left\| \frac{1}{T} \sum_{t=1}^{T} \mathbf{F}_t \mathbf{H}_2^\top \mathbf{H}_2 \mathbf{F}_t^\top \mathbf{H}_1 \right\|_2$$
$$\geq \sqrt{p_1} p_2 \left\| \frac{1}{T} \sum_{t=1}^{T} \mathbf{F}_t \mathbf{F}_t^\top \right\|_2 \nu_{\min}(\mathbf{H}_1) \nu_{\min}^2(\mathbf{H}_2) \gtrsim \sqrt{p_1} p_2 \nu_{\min}(\mathbf{H}_1) \nu_{\min}^2(\mathbf{H}_2).$$

因此，我们有

$$\|\boldsymbol{\delta}_1\|_F \asymp O_p\left(\sqrt{p_1} p_2 \nu_{\min}(\mathbf{H}_1) \nu_{\min}^2(\mathbf{H}_2)\right).$$

对于第二项 $\boldsymbol{\delta}_2$，我们有

$$\|\boldsymbol{\delta}_2\|_F = \left\| \frac{1}{T p_1 p_2} \sum_{t=1}^{T} \mathbf{R}\mathbf{F}_t \mathbf{C}^\top \mathbf{W}_2 \mathbf{W}_2^\top \mathbf{E}_t^\top \mathbf{W}_1 \right\|_F \leq \left\| \frac{1}{T p_1 p_2} \sum_{t=1}^{T} \mathbf{R}\mathbf{F}_t \mathbf{W}_2^\top \mathbf{E}_t^\top \mathbf{W}_1 \right\|_F \|\mathbf{C}^\top \mathbf{W}_2\|_2$$
$$\leq \frac{1}{T p_1} \|\mathbf{R}\|_F \| \sum_{t=1}^{T} \mathbf{F}_t \mathbf{W}_2^\top \mathbf{E}_t^\top \mathbf{W}_1 \|_F \|\mathbf{H}_2\|_2 = O_p\left(\sqrt{\frac{p_2}{T}} \nu_{\min}(\mathbf{H}_2) \right),$$

其中，最后一个等式根据引理 S2.3 (3) 得出。

根据引理 S2.3 (3)，我们有 $\mathbb{E}\| \sum_{t=1}^{T} \mathbf{F}_t \mathbf{W}_2^\top \mathbf{E}_t^\top \|_F^2 = O(T p_1 p_2)$，因此，

$$\|\boldsymbol{\delta}_3\|_F = \left\| \frac{1}{T p_1 p_2} \sum_{t=1}^{T} \mathbf{E}_t \mathbf{W}_2 \mathbf{W}_2^\top \mathbf{C} \mathbf{F}_t^\top \mathbf{R}^\top \mathbf{W}_1 \right\|_F \leq \frac{1}{T p_1 p_2} \| \sum_{t=1}^{T} \mathbf{E}_t \mathbf{W}_2 \mathbf{F}_t^\top \|_F \|\mathbf{W}_2^\top \mathbf{C}\|_2 \|\mathbf{R}^\top \mathbf{W}_1\|_2$$
$$= O_p\left(\sqrt{\frac{p_1 p_2}{T}} \nu_{\min}(\mathbf{H}_1) \nu_{\min}(\mathbf{H}_2) \right).$$

根据引理 S2.3 (5)，

$$\|\boldsymbol{\delta}_4\|_F = \left\| \frac{1}{T p_1 p_2} \sum_{t=1}^{T} \mathbf{E}_t \mathbf{W}_2 \mathbf{W}_2^\top \mathbf{E}_t^\top \mathbf{W}_1 \right\|_F = O_p\left(\sqrt{\frac{p_2}{T}} + \frac{1}{\sqrt{p_1}} \right).$$

令 $\mathbf{Z} = \frac{1}{T^2} \left(\sum_{t=1}^{T} \widehat{\mathbf{F}}_t \mathbf{W}_2^\top \mathbf{X}_t^\top \right) \left(\sum_{t=1}^{T} \mathbf{X}_t \mathbf{W}_2 \widehat{\mathbf{F}}_t^\top \right)$，则

$$\mathbf{Z} = \frac{1}{p_1^2 p_2^2 T^2} \left(\sum_{t=1}^{T} \mathbf{W}_1^\top \mathbf{X}_t \mathbf{W}_2 \mathbf{W}_2^\top \mathbf{X}_t^\top \right) \left(\sum_{t=1}^{T} \mathbf{X}_t \mathbf{W}_2 \mathbf{W}_2^\top \mathbf{X}_t^\top \mathbf{W}_1 \right)$$
$$= (\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2 + \boldsymbol{\delta}_3 + \boldsymbol{\delta}_4)^\top (\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2 + \boldsymbol{\delta}_3 + \boldsymbol{\delta}_4) = \sum_{i=1}^{4} \sum_{j=1}^{4} \boldsymbol{\delta}_i^\top \boldsymbol{\delta}_j.$$

我们可以证明，当 $\nu_{\min}^2(\mathbf{H}_2) \gg \max(\frac{1}{T}, \frac{1}{p_2})$ 时，$\|\mathbf{Z}\|_F \geq \|\mathbf{Z}\|_2 \gtrsim \|\boldsymbol{\delta}_1\|_2^2 \gtrsim p_1 p_2^2 \nu_{\min}^2(\mathbf{H}_1) \nu_{\min}^4(\mathbf{H}_2)$，因此

$$\|\mathbf{Z}^{-1/2}\|_F = O_p\left(\frac{1}{\sqrt{p_1} p_2 \nu_{\min}(\mathbf{H}_1) \nu_{\min}^2(\mathbf{H}_2)} \right).$$

$$
\begin{aligned}
\widehat{\mathbf{R}}^{(1)} &= \sqrt{p_1}\left(\sum_{t=1}^{T}\mathbf{X}_t\mathbf{W}_2\widehat{\mathbf{F}}_t^\top\right)\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t\mathbf{W}_2^\top\mathbf{X}_t^\top\right)\left(\sum_{t=1}^{T}\mathbf{X}_t\mathbf{W}_2\widehat{\mathbf{F}}_t^\top\right)\right]^{-1/2}\\
&= \sqrt{p_1}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{X}_t\mathbf{W}_2\widehat{\mathbf{F}}_t^\top\right)\mathbf{Z}^{-1/2}\\
&= \mathbf{R}\left[\frac{1}{\sqrt{p_1}p_2}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\right)\mathbf{Z}^{-1/2}\right] + \frac{1}{\sqrt{p_1}p_2}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\right)\mathbf{Z}^{-1/2}\\
&+ \frac{1}{\sqrt{p_1}p_2}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\right)\mathbf{Z}^{-1/2} + \frac{1}{\sqrt{p_1}p_2}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\right)\mathbf{Z}^{-1/2}\\
&:= \mathcal{I}+\mathcal{II}+\mathcal{III}+\mathcal{IV}.
\end{aligned}
$$

令 $\widehat{\mathrm{H}}_r^{(1)} = \frac{1}{\sqrt{p_1}p_2T}\left(\sum_{t=1}^{T}\mathrm{F}_t\mathrm{C}^\top\mathrm{W}_2\mathrm{W}_2^\top\mathrm{C}\mathrm{F}_t^\top\mathrm{R}^\top\mathrm{W}_1\right)\mathrm{Z}^{-1/2}$，则，

$$
\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)} = \mathcal{II}+\mathcal{III}+\mathcal{IV}. \tag{S1.1}
$$

As

$$
\begin{aligned}
&\left\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\right\|_F \le \left\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{F}_t^\top\right\|_F \|\mathbf{C}^\top\mathbf{W}_2\|_2\|\mathbf{W}_2^\top\mathbf{C}\|_2\|\mathbf{R}^\top\mathbf{W}_1\|_2\\
&= p_1p_2^2\left\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t^\top\mathbf{F}_t\right\|_2\|\mathbf{H}_2\|_2^2\|\mathbf{H}_1\|_2 = O_p\left(p_1p_2^2\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)\right),
\end{aligned}
$$

因此，我们得到以下结果:

$$
\begin{aligned}
\|\widehat{\mathbf{H}}_r^{(1)}\|_F &= \left\|\frac{1}{\sqrt{p_1}p_2}(\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}\right\|_F\\
&\le \frac{1}{\sqrt{p_1}p_2}\left\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\right\|_F\|\mathbf{Z}^{-1/2}\|_F = O_p(1);
\end{aligned}
$$

$$
\|\mathcal{II}\|_F = \left\|\frac{1}{\sqrt{p_1}p_2}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\right)\mathbf{Z}^{-1/2}\right\|_F = O_p\left(\frac{1}{\sqrt{Tp_2}}\nu_{\min}^{-1}(\mathbf{H}_1)\nu_{\min}^{-1}(\mathbf{H}_2)\right);
$$

$$
\|\mathcal{III}\|_F = \left\|\frac{1}{\sqrt{p_1}p_2}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\right)\mathbf{Z}^{-1/2}\right\|_F = O_p\left(\sqrt{\frac{p_1}{Tp_2}}\nu_{\min}^{-1}(\mathbf{H}_2)\right);
$$

$$
\|\mathcal{IV}\|_F = \left\|\frac{1}{\sqrt{p_1}p_2}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\right)\mathbf{Z}^{-1/2}\right\|_F = O_p\left(\frac{1}{\sqrt{Tp_2}}\nu_{\min}^{-1}(\mathbf{H}_1)\nu_{\min}^{-2}(\mathbf{H}_2)+\frac{1}{\sqrt{p_1}p_2}\nu_{\min}^{-1}(\mathbf{H}_1)\nu_{\min}^{-2}(\mathbf{H}_2)\right);
$$

最后，我们可以得到

$$
\frac{1}{p_1}\|\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\|_F^2 = O_p\left(\frac{1}{Tp_2\nu_{\min}^2(\mathbf{H}_2)}+\frac{1}{Tp_1p_2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}+\frac{1}{p_1^2p_2^2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}\right).
$$

还需证明 $\widehat{\mathbf{H}}_r^{(1)\top}\widehat{\mathbf{H}}_r^{(1)} \xrightarrow{p} \mathrm{I}_{k_1}$。在满足条件 $\nu_{\min}^2(\mathbf{H}_2) \gg \max(\frac{1}{T}, \frac{1}{p_2})$ 时，我们可以得到

$$\frac{1}{p_1}\|\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\|_F^2 = o_p(1),$$

$$\|\frac{1}{p_1}\mathbf{R}^\top(\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\|_F \le (\frac{\|\mathbf{R}\|_F^2}{p_1}\frac{\|\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\|_F^2}{p_1})^{1/2} = o_p(1),\ \|\frac{1}{p_1}\widehat{\mathbf{R}}^{(1)\top}(\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\|_F = o_p(1).$$

注意，$p_1^{-1}\widehat{\mathbf{R}}^{(1)\top}\widehat{\mathbf{R}}^{(1)} = \mathrm{I}_{k_1}$，而 $p_1^{-1}\mathbf{R}^\top\mathbf{R} = \mathrm{I}_{k_1}$，于是

$$\mathbf{I}_{k_1} = \frac{1}{p_1}\widehat{\mathbf{R}}^{(1)\top}\mathbf{R}\widehat{\mathbf{H}}_r^{(1)} + o_p(1) = \widehat{\mathbf{H}}_r^{(1)\top}\widehat{\mathbf{H}}_r^{(1)} + o_p(1).$$

下面我们证明 $\widehat{\mathbf{R}}^{(1)}$ 的逐行一致性。由公式（S1.1），可得

$$\widehat{\boldsymbol{R}}_{i\cdot}^{(1)} - \widehat{\mathbf{H}}_r^{(1)\top}\boldsymbol{R}_{i\cdot} = \frac{1}{T\sqrt{p_1}p_2}\mathbf{Z}^{-1/2}\left(\sum_{t=1}^T \mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\boldsymbol{R}_{i\cdot}\right) + \frac{1}{T\sqrt{p_1}p_2}\mathbf{Z}^{-1/2}\left(\sum_{t=1}^T \mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\boldsymbol{e}_{t,i\cdot}\right)$$
$$+ \frac{1}{T\sqrt{p_1}p_2}\mathbf{Z}^{-1/2}\left(\sum_{t=1}^T \mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\boldsymbol{e}_{t,i\cdot}\right).$$

由引理 S2.3 (3)，可得

$$\left\|\frac{1}{T\sqrt{p_1}p_2}\mathbf{Z}^{-1/2}\left(\sum_{t=1}^T \mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\boldsymbol{R}_{i\cdot}\right)\right\|_2 \le \frac{1}{T\sqrt{p_1}p_2}\left\|\mathbf{Z}^{-1/2}\right\|_F\left\|\sum_{t=1}^T \mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\mathbf{F}_t^\top\right\|_F\left\|\mathbf{W}_2^\top\mathbf{C}\right\|_2$$
$$= O_p\left(\frac{1}{\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}\right).$$

与引理S2.3 (3)的证明类似，可得 $\left\|\sum_{t=1}^T \mathrm{F}_t\mathrm{W}_2^\top\boldsymbol{e}_{t,\ i\cdot}\right\|_F^2 = O_p(Tp_2)$，于是

$$\left\|\frac{1}{T\sqrt{p_1}p_2}\mathbf{Z}^{-1/2}\left(\sum_{t=1}^T \mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\boldsymbol{e}_{t,i\cdot}\right)\right\|_2 \le \frac{1}{T\sqrt{p_1}p_2}\left\|\mathbf{Z}^{-1/2}\right\|_F\left\|\mathbf{W}_1^\top\mathbf{R}\right\|_2\left\|\mathbf{C}^\top\mathbf{W}_2\right\|_2\left\|\sum_{t=1}^T \mathbf{F}_t\mathbf{W}_2^\top\boldsymbol{e}_{t,i\cdot}\right\|_F$$
$$= O_p\left(\frac{1}{\sqrt{Tp_2}\nu_{\min}(\mathbf{H}_2)}\right).$$

与引理S2.3 (5)的证明类似，也可得 $\left\|\sum_{t=1}^T \mathrm{W}_1^\top\mathrm{E}_t\mathrm{W}_2\mathrm{W}_2^\top\boldsymbol{e}_{t,\ i\cdot}\right\|_F^2 = O_p\left(Tp_1p_2^3 + T^2p_2^2\right)$,

$$\left\|\frac{1}{T\sqrt{p_1}p_2}\mathbf{Z}^{-1/2}\left(\sum_{t=1}^T \mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\boldsymbol{e}_{t,i\cdot}\right)\right\|_2 \le \frac{1}{T\sqrt{p_1}p_2}\left\|\mathbf{Z}^{-1/2}\right\|_F\left\|\sum_{t=1}^T \mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\boldsymbol{e}_{t,i\cdot}\right\|_F$$
$$= O_p\left(\frac{1}{\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)} + \frac{1}{p_1p_2\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}\right).$$

结合上述结果，对于任意 $i \le p_1$，我们可以得到

$$\|\widehat{\boldsymbol{R}}_{i\cdot}^{(1)} - \widehat{\mathbf{H}}_r^{(1)\top}\boldsymbol{R}_{i\cdot}\|_2^2 = O_p\left(\frac{1}{Tp_2\nu_{\min}^2(\mathbf{H}_2)} + \frac{1}{Tp_1p_2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)} + \frac{1}{p_1^2p_2^2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}\right).$$

□

S1.4 定理3.2 的证明

证明。首先，我们将 $\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(1)}\widehat{\mathbf{F}}_t$ 分解为四项：

$$\begin{aligned}\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(1)}\widehat{\mathbf{F}}_t &= \frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{X}_t\mathbf{W}_2\\ &= \frac{1}{p_1p_2}\sum_{t=1}^{T}(\mathbf{R}\mathbf{F}_t\mathbf{C}^\top+\mathbf{E}_t)^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top(\mathbf{R}\mathbf{F}_t\mathbf{C}^\top+\mathbf{E}_t)\mathbf{W}_2\\ &= \frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2+\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\\ &+\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2+\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\\ &= \boldsymbol{\delta}_1^{(1)}+\boldsymbol{\delta}_2^{(1)}+\boldsymbol{\delta}_3^{(1)}+\boldsymbol{\delta}_4^{(1)}.\end{aligned}$$

对于第一项，

$$\begin{aligned}\left\|\frac{1}{T}\boldsymbol{\delta}_1^{(1)}\right\|_F &= \left\|\frac{1}{Tp_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\right\|_F \asymp \left\|\frac{1}{Tp_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\right\|_F\\ &= p_1\|\mathbf{C}\|_F\left\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t^\top\mathbf{F}_t\right\|_F\|\mathbf{H}_1\|_2\|\mathbf{H}_2\|_2 = O_p\left(p_1\sqrt{p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)\right).\end{aligned}$$

此外，我们还有 $\|\boldsymbol{\delta}_1^{(1)}\|_F \geq \|\boldsymbol{\delta}_1^{(1)}\|_2 \gtrsim Tp_1\sqrt{p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)$，因此，我们有

$$\|\boldsymbol{\delta}_1^{(1)}\|_F \asymp O_p\left(Tp_1\sqrt{p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)\right).$$

由引理 S2.4，可得

$$\left\|\sum_{s=1}^{T}\mathbf{E}_s^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{F}_s\right\|_F^2 = O_p\left(\frac{p_1^2}{p_2\nu_{\min}^2(\mathbf{H}_2)}+\frac{p_1}{\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}+\frac{T}{p_2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}\right),$$

then

$$\begin{aligned}\|\boldsymbol{\delta}_2^{(1)}\|_F &= \left\|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\right\|_F = \left\|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}+\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\right\|_F\\ &\leq \left\|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\right\|_F+\left\|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\right\|_F\\ &\leq \left\|\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{F}_t\right\|_F\|\mathbf{H}_1\|_2\|\mathbf{H}_2\|_2+\left\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\mathbf{F}_t\right\|_F\|\mathbf{H}_1\|_2\|\mathbf{H}_2\|_2\left\|\widehat{\mathbf{H}}_r^{(1)}\right\|_F\\ &= O_p\left(\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)+\frac{p_1}{\sqrt{p_2}}\nu_{\min}(\mathbf{H}_1)+\frac{\sqrt{p_1}}{\nu_{\min}(\mathbf{H}_2)}\right).\end{aligned}$$

与引理的证明S2.3 (3)类似，我们也有 $\mathbb{E}\|\sum_{t=1}^{T}\mathbf{F}_t^\top\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\|_F^2 = O\left(Tp_1p_2\right)$，因此

$$\left\|\boldsymbol{\delta}_3^{(1)}\right\|_F = \left\|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F \asymp \frac{1}{p_2}\left\|\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F$$
$$\leq \frac{1}{p_2}\|\mathbf{C}\|_F\left\|\sum_{t=1}^{T}\mathbf{F}_t^\top\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F = O_p\left(\sqrt{Tp_1}\right).$$

在引理的证明S2.3(5)中，我们已证明 $\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\|_F^2 = O_p\left(Tp_1^2p_2^2+T^2p_1p_2\right)$，因此

$$\left\|\boldsymbol{\delta}_4^{(1)}\right\|_F = \left\|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F = \left\|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}+\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F$$
$$\leq \left\|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F + \left\|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F$$
$$\leq \frac{1}{p_1p_2}\left\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F\left\|\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\right\|_F + \frac{1}{p_1p_2}\left\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F\|\mathbf{R}\|_F\|\widehat{\mathbf{H}}_r^{(1)}\|_F$$
$$= O_p\left(\sqrt{Tp_1}+\frac{T}{\sqrt{p_2}}\right).$$

令 $\mathrm{Z}^{(1)} = (\sum_{t=1}^{T}\widehat{\mathrm{F}}_t^\top\widehat{\mathrm{R}}^{(1)\top}\mathrm{X}_t)(\sum_{t=1}^{T}\mathrm{X}_t^\top\widehat{\mathrm{R}}^{(1)}\widehat{\mathrm{F}}_t)$，则 $\mathrm{Z}^{(1)} = \sum_{i=1}^{4}\sum_{j=1}^{4}\boldsymbol{\delta}_i^{(1)\top}\boldsymbol{\delta}_j^{(1)}$。我们可以证明$\|\mathrm{Z}^{(1)}\|_F \gtrsim T^2p_1^2p_2\nu_{\min}^2(\mathrm{H}_1)\nu_{\min}^2(\mathrm{H}_2)$，因此 $\|(\mathrm{Z}^{(1)})^{-1/2}\|_F = O_p\left(\frac{1}{Tp_1\sqrt{p_2}\nu_{\min}(\mathrm{H}_1)\nu_{\min}(\mathrm{H}_2)}\right)$。

$$\widehat{\mathbf{C}}^{(1)} = \sqrt{p_2}\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(1)}\widehat{\mathbf{F}}_t\right)\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^\top\widehat{\mathbf{R}}^{(1)\top}\mathbf{X}_t\right)\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(1)}\widehat{\mathbf{F}}_t\right)\right]^{-1/2}$$
$$= \sqrt{p_2}\left(\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\right)\left(\mathbf{Z}^{(1)}\right)^{-1/2}$$
$$+ \sqrt{p_2}\left(\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\right)\left(\mathbf{Z}^{(1)}\right)^{-1/2}$$
$$+ \sqrt{p_2}\left(\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right)\left(\mathbf{Z}^{(1)}\right)^{-1/2}$$
$$+ \sqrt{p_2}\left(\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right)\left(\mathbf{Z}^{(1)}\right)^{-1/2}.$$
$$= \sqrt{p_2}\left(\boldsymbol{\delta}_1^{(1)}(\mathbf{Z}^{(1)})^{-1/2}+\boldsymbol{\delta}_2^{(1)}(\mathbf{Z}^{(1)})^{-1/2}+\boldsymbol{\delta}_3^{(1)}(\mathbf{Z}^{(1)})^{-1/2}+\boldsymbol{\delta}_4^{(1)}(\mathbf{Z}^{(1)})^{-1/2}\right).$$

令 $\widehat{\mathrm{H}}_c^{(1)} = \left(\frac{1}{p_1\sqrt{p_2}}\sum_{t=1}^{T}\mathrm{F}_t^\top\mathrm{R}^\top\widehat{\mathrm{R}}^{(1)}\mathrm{W}_1^\top\mathrm{R}\mathrm{F}_t\mathrm{C}^\top\mathrm{W}_2\right)\left(\mathrm{Z}^{(1)}\right)^{-1/2}$，则

$$\widehat{\mathbf{C}}^{(1)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(1)} = \sqrt{p_2}\boldsymbol{\delta}_2^{(1)}(\mathbf{Z}^{(1)})^{-1/2} + \sqrt{p_2}\boldsymbol{\delta}_3^{(1)}(\mathbf{Z}^{(1)})^{-1/2} + \sqrt{p_2}\boldsymbol{\delta}_4^{(1)}(\mathbf{Z}^{(1)})^{-1/2}.$$

$$\begin{aligned}\|\widehat{\mathbf{H}}_c^{(1)}\|_F &= \left\| \left(\frac{1}{p_1\sqrt{p_2}} \sum_{t=1}^{T} \mathbf{F}_t^\top \mathbf{R}^\top \widehat{\mathbf{R}}^{(1)} \mathbf{W}_1^\top \mathbf{R} \mathbf{F}_t \mathbf{C}^\top \mathbf{W}_2 \right) \left(\mathbf{Z}^{(1)} \right)^{-1/2} \right\|_F \\ &\leq \left\| \frac{1}{p_1\sqrt{p_2}} \sum_{t=1}^{T} \mathbf{F}_t^\top \mathbf{R}^\top \widehat{\mathbf{R}}^{(1)} \mathbf{W}_1^\top \mathbf{R} \mathbf{F}_t \mathbf{C}^\top \mathbf{W}_2 \right\|_F \left\| \left(\mathbf{Z}^{(1)} \right)^{-1/2} \right\|_F \\ &= O_p(1).\end{aligned}$$

$$\|\boldsymbol{\delta}_2^{(1)} (\mathbf{Z}^{(1)})^{-1/2}\|_F = O_p\left(\frac{1}{\sqrt{Tp_1}} + \frac{1}{Tp_2} \nu_{\min}^{-1}(\mathbf{H}_2) + \frac{1}{T\sqrt{p_1 p_2}} \nu_{\min}^{-1}(\mathbf{H}_1) \nu_{\min}^{-2}(\mathbf{H}_2) \right).$$

$$\|\boldsymbol{\delta}_3^{(1)} (\mathbf{Z}^{(1)})^{-1/2}\|_F = O_p\left(\frac{1}{\sqrt{Tp_1p_2}} \nu_{\min}^{-1}(\mathbf{H}_1) \nu_{\min}^{-1}(\mathbf{H}_2) \right).$$

$$\|\boldsymbol{\delta}_4^{(1)} (\mathbf{Z}^{(1)})^{-1/2}\|_F = O_p\left(\frac{1}{\sqrt{Tp_1p_2}} \nu_{\min}^{-1}(\mathbf{H}_1) \nu_{\min}^{-1}(\mathbf{H}_2) + \frac{1}{p_1p_2} \nu_{\min}^{-1}(\mathbf{H}_1) \nu_{\min}^{-1}(\mathbf{H}_2) \right).$$

因此

$$\frac{1}{p_2}\|\widehat{\mathbf{C}}^{(1)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(1)}\|_F^2 = O_p\left(\frac{1}{Tp_1} + \frac{1}{p_1^2 p_2^2 \nu_{\min}^2(\mathbf{H}_1) \nu_{\min}^2(\mathbf{H}_2)} + \frac{1}{T^2 p_2^2 \nu_{\min}^2(\mathbf{H}_2)} + \frac{1}{Tp_1p_2 \nu_{\min}^2(\mathbf{H}_1) \nu_{\min}^2(\mathbf{H}_2)} \right).$$

$$\left\| \frac{1}{p_2} \mathbf{C}^\top (\widehat{\mathbf{C}}^{(1)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(1)}) \right\|_F \leq \left(\frac{\|\mathbf{C}\|_F^2}{p_2} \frac{\|\widehat{\mathbf{C}}^{(1)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(1)}\|_F^2}{p_2} \right)^{1/2} = o_p(1), \left\| \frac{1}{p_2} \widehat{\mathbf{C}}^{(1)\top} (\widehat{\mathbf{C}}^{(1)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(1)}) \right\|_F = o_p(1).$$

注意到 $p_2^{-1}\mathbf{C}^\top\mathbf{C} = \mathbf{I}_{k_2}$, $p_2^{-1}\widehat{\mathbf{C}}^{(1)\top}\widehat{\mathbf{C}}^{(1)} = \mathbf{I}_{k_2}$, 于是

$$\mathbf{I}_{k_2} = \frac{1}{p_2} \widehat{\mathbf{C}}^{(1)\top} \mathbf{C} \widehat{\mathbf{H}}_c^{(1)} + o_p(1) = \widehat{\mathbf{H}}_c^{(1)\top} \widehat{\mathbf{H}}_c^{(1)} + o_p(1), \ \widehat{\mathbf{H}}_c^{(1)\top} \widehat{\mathbf{H}}_c^{(1)} \xrightarrow{p} \mathbf{I}_{k_2}.$$

接下来，我们证明 $\widehat{\mathbf{C}}^{(1)}$ 的逐行一致性。

对于 $j \leq p_2$,

$$\begin{aligned}\widehat{\boldsymbol{C}}_{j\cdot}^{(1)} - \widehat{\mathbf{H}}_c^{(1)\top} \boldsymbol{C}_{j\cdot} &= \frac{1}{p_1\sqrt{p_2}} (\mathbf{Z}^{(1)})^{-1/2} \left(\sum_{t=1}^{T} \mathbf{W}_2^\top \mathbf{C} \mathbf{F}_t^\top \mathbf{R}^\top \mathbf{W}_1 \widehat{\mathbf{R}}^{(1)\top} \boldsymbol{e}_{t,\cdot j} \right) \\ &+ \frac{1}{p_1\sqrt{p_2}} (\mathbf{Z}^{(1)})^{-1/2} \left(\sum_{t=1}^{T} \mathbf{W}_2^\top \mathbf{E}_t^\top \mathbf{W}_1 \widehat{\mathbf{R}}^{(1)\top} \mathbf{R} \mathbf{F}_t \boldsymbol{C}_{j\cdot} \right) \\ &+ \frac{1}{p_1\sqrt{p_2}} (\mathbf{Z}^{(1)})^{-1/2} \left(\sum_{t=1}^{T} \mathbf{W}_2^\top \mathbf{E}_t^\top \mathbf{W}_1 \widehat{\mathbf{R}}^{(1)\top} \boldsymbol{e}_{t,\cdot j} \right).\end{aligned}$$

首先，类似于引理的证明S2.4，我们可以得到

$$\left\| \sum_{t=1}^{T} \mathbf{F}_t^\top (\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)})^\top \boldsymbol{e}_{t,\cdot j} \right\|_F^2 = O_p\left(\frac{p_1^2}{p_2^2 \nu_{\min}^2(\mathbf{H}_2)} + \frac{p_1}{p_2 \nu_{\min}^2(\mathbf{H}_1) \nu_{\min}^4(\mathbf{H}_2)} + \frac{T}{p_2^2 \nu_{\min}^2(\mathbf{H}_1) \nu_{\min}^4(\mathbf{H}_2)} \right),$$

$$\begin{aligned}\left\|\frac{1}{p_1\sqrt{p_2}}\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\widehat{\mathbf{R}}^{(1)\top}\boldsymbol{e}_{t,\cdot j}\right\|_2 &\le \frac{1}{p_1\sqrt{p_2}}\left\|\sum_{t=1}^{T}\mathbf{F}_t^\top\widehat{\mathbf{R}}^{(1)\top}\boldsymbol{e}_{t,\cdot j}\right\|_F\|\mathbf{W}_2^\top\mathbf{C}\|_2\|\mathbf{R}^\top\mathbf{W}_1\|_2\\ &\le \frac{1}{p_1\sqrt{p_2}}\left\|\sum_{t=1}^{T}\mathbf{F}_t^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}_t\widehat{\mathbf{H}}_r^{(1)})^\top\boldsymbol{e}_{t,\cdot j}\right\|_F\|\mathbf{W}_2^\top\mathbf{C}\|_2\|\mathbf{R}^\top\mathbf{W}_1\|_2\\ &+\frac{1}{p_1\sqrt{p_2}}\left\|\sum_{t=1}^{T}\mathbf{F}_t^\top\mathbf{R}_t^\top\boldsymbol{e}_{t,\cdot j}\right\|_F\|\mathbf{W}_2^\top\mathbf{C}\|_2\|\mathbf{R}^\top\mathbf{W}_1\|_2\|\widehat{\mathbf{H}}_c^{(1)}\|_F,\end{aligned}$$

然后,

$$\begin{aligned}\left\|\frac{1}{p_1\sqrt{p_2}}(\mathbf{Z}^{(1)})^{-1/2}\left(\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\widehat{\mathbf{R}}^{(1)\top}\boldsymbol{e}_{t,\cdot j}\right)\right\|_2 &\le \frac{1}{p_1\sqrt{p_2}}\left\|\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\widehat{\mathbf{R}}^{(1)\top}\boldsymbol{e}_{t,\cdot j}\right\|_F\left\|(\mathbf{Z}^{(1)})^{-1/2}\right\|_F\\ &= O_p\left(\frac{1}{Tp_2\nu_{\min}(\mathbf{H}_2)}+\frac{1}{T\sqrt{p_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}+\frac{1}{\sqrt{Tp_1}}\right).\end{aligned}$$

其次,

$$\begin{aligned}\left\|\frac{1}{p_1\sqrt{p_2}}(\mathbf{Z}^{(1)})^{-1/2}\left(\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\widehat{\mathbf{R}}^{(1)\top}\mathbf{R}\mathbf{F}_t\boldsymbol{C}_{j\cdot}\right)\right\|_2 &\asymp \left\|\frac{1}{\sqrt{p_2}}(\mathbf{Z}^{(1)})^{-1/2}\left(\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\mathbf{F}_t\right)\right\|_2\\ &= O_p\left(\frac{1}{\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}\right).\end{aligned}$$

第三,

$$\begin{aligned}\left\|\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\widehat{\mathbf{R}}^{(1)\top}\boldsymbol{e}_{t,\cdot j}\right\|_F &= \left\|\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}+\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})^\top\boldsymbol{e}_{t,\cdot j}\right\|_F\\ &\le \left\|\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\boldsymbol{e}_{t,\cdot j}^\top\right\|_F\left\|\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\right\|_F\\ &+\left\|\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\boldsymbol{e}_{t,\cdot j}^\top\right\|_F\|\mathbf{R}\|_F\|\widehat{\mathbf{H}}_r^{(1)\top}\|_F,\end{aligned}$$

$$\begin{aligned}\left\|\frac{1}{p_1\sqrt{p_2}}(\mathbf{Z}^{(1)})^{-1/2}\left(\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\widehat{\mathbf{R}}^{(1)\top}\boldsymbol{e}_{t,\cdot j}\right)\right\|_F &\le \frac{1}{p_1\sqrt{p_2}}\left\|(\mathbf{Z}^{(1)})^{-1/2}\left(\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\widehat{\mathbf{R}}^{(1)\top}\boldsymbol{e}_{t,\cdot j}\right)\right\|_F\\ &= O_p\left(\frac{1}{p_1p_2\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}+\frac{1}{\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}\right).\end{aligned}$$

因此，我们得到

$$\|\widehat{\boldsymbol{C}}_{j\cdot}^{(1)}-\widehat{\mathbf{H}}_c^{(1)\top}\boldsymbol{C}_{j\cdot}\|_2^2=O_p\left(\frac{1}{Tp_1}+\frac{1}{p_1^2p_2^2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}+\frac{1}{T^2p_2^2\nu_{\min}^2(\mathbf{H}_2)}+\frac{1}{Tp_1p_2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}\right).$$

□

S1.5 定理的证明 3.3

证明。根据定义，

$$\begin{aligned}\widehat{\mathbf{F}}_t^{(2)} &= \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(1)\top}\mathbf{X}_t\widehat{\mathbf{C}}^{(1)} = \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(1)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(1)} + \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(1)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(1)} \\ &= \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(1)\top}\left(\mathbf{R}-\widehat{\mathbf{R}}^{(1)}(\widehat{\mathbf{H}}_r^{(1)})^{-1}+\widehat{\mathbf{R}}^{(1)}(\widehat{\mathbf{H}}_r^{(1)})^{-1}\right)\mathbf{F}_t\left(\mathbf{C}-\widehat{\mathbf{C}}^{(1)}(\widehat{\mathbf{H}}_c^{(1)})^{-1}+\widehat{\mathbf{C}}^{(1)}(\widehat{\mathbf{H}}_c^{(1)})^{-1}\right)^\top\widehat{\mathbf{C}}^{(1)} \\ &+ \frac{1}{p_1p_2}\left(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}+\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\right)^\top\mathbf{E}_t\left(\widehat{\mathbf{C}}^{(1)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(1)}+\mathbf{C}\widehat{\mathbf{H}}_c^{(1)}\right).\end{aligned}$$

注意，$\widehat{\mathbf{R}}^{(1)\top}\widehat{\mathbf{R}}^{(1)} = p_1\mathbf{I}_{k_1}$，$\widehat{\mathbf{C}}^{(1)\top}\widehat{\mathbf{C}}^{(1)} = p_2\mathbf{I}_{k_2}$，则

$$\begin{aligned}\widehat{\mathbf{F}}_t^{(2)} - (\widehat{\mathbf{H}}_r^{(1)})^{-1}\mathbf{F}_t\left((\widehat{\mathbf{H}}_c^{(1)})^{-1}\right)^\top &= \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(1)\top}\left(\mathbf{R}-\widehat{\mathbf{R}}^{(1)}(\widehat{\mathbf{H}}_r^{(1)})^{-1}\right)\mathbf{F}_t\left(\mathbf{C}-\widehat{\mathbf{C}}(\widehat{\mathbf{H}}_c^{(1)})^{-1}\right)^\top\widehat{\mathbf{C}}^{(1)} \\ &+ \frac{1}{p_1}\widehat{\mathbf{R}}^{(1)\top}\left(\mathbf{R}-\widehat{\mathbf{R}}^{(1)}(\widehat{\mathbf{H}}_r^{(1)})^{-1}\right)\mathbf{F}_t\left((\widehat{\mathbf{H}}_c^{(1)})^{-1}\right)^\top \\ &+ \frac{1}{p_2}(\widehat{\mathbf{H}}_r^{(1)})^{-1}\mathbf{F}_t\left(\mathbf{C}-\widehat{\mathbf{C}}(\widehat{\mathbf{H}}_c^{(1)})^{-1}\right)^\top\widehat{\mathbf{C}}^{(1)} \\ &+ \frac{1}{p_1p_2}\left(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\right)^\top\mathbf{E}_t\left(\widehat{\mathbf{C}}^{(1)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(1)}\right) \\ &+ \frac{1}{p_1p_2}\left(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\right)^\top\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(1)} + \frac{1}{p_1p_2}\widehat{\mathbf{H}}_r^{(1)\top}\mathbf{R}^\top\mathbf{E}_t\left(\widehat{\mathbf{C}}^{(1)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(1)}\right) \\ &+ \frac{1}{p_1p_2}\widehat{\mathbf{H}}_r^{(1)\top}\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(1)}.\end{aligned}$$

在引理 S2.5中，我们证明了

$$\left\|\mathbf{R}^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\right\|_F = O_p\left(\frac{\sqrt{p_1}}{\sqrt{Tp_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}+\frac{1}{p_2\nu_{\min}(\mathbf{H}_1)\nu^2_{\min}(\mathbf{H}_2)}\right),$$

$$\left\|\mathbf{C}^\top(\widehat{\mathbf{C}}^{(1)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(1)})\right\|_F = O_p\left(\frac{1}{T\nu_{\min}(\mathbf{H}_2)}+\frac{\sqrt{p_2}}{\sqrt{Tp_1}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}+\frac{1}{p_1\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}\right).$$

因此，根据柯西-施瓦茨不等式、定理 3.1、定理 3.2以及 $\|\mathbf{R}^\top\mathbf{E}_t\|_F^2$、$\|\mathbf{E}_t\mathbf{C}\|_F^2$ 和 $\|\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\|_F^2$的界，我们有

$$\left\|\widehat{\mathbf{F}}_t^{(2)} - (\widehat{\mathbf{H}}_r^{(1)})^{-1}\mathbf{F}_t\left((\widehat{\mathbf{H}}_c^{(1)})^{-1}\right)^\top\right\|_F = O_p\left(\frac{1}{\sqrt{T}p_1}+\frac{1}{\sqrt{T}p_2\nu_{\min}(\mathbf{H}_2)}+\frac{1}{\sqrt{p_1p_2}}+\gamma_f\right),$$

其中

$$\gamma_f = \frac{1}{\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}+\frac{1}{p_1p_2\nu_{\min}(\mathbf{H}_1)\nu^2_{\min}(\mathbf{H}_2)}.$$

接下来，对于任意 t,i,j

$$\begin{aligned}\widehat{S}_{t,ij}^{(2)} - S_{t,ij} &= \widehat{\boldsymbol{R}}_{i\cdot}^{(1)\top}\widehat{\mathbf{F}}_t^{(2)}\widehat{\boldsymbol{C}}_{j\cdot}^{(1)} - \boldsymbol{R}_{i\cdot}^\top\mathbf{F}_t\boldsymbol{C}_{j\cdot} \\ &= (\widehat{\boldsymbol{R}}_{i\cdot}^{(1)} - \widehat{\mathbf{H}}_r^{(1)\top}\boldsymbol{R}_{i\cdot} + \widehat{\mathbf{H}}_r^{(1)\top}\boldsymbol{R}_{i\cdot})^\top\widehat{\mathbf{F}}_t^{(2)}(\widehat{\boldsymbol{C}}_{j\cdot}^{(1)} - \widehat{\mathbf{H}}_c^{(1)\top}\boldsymbol{C}_{j\cdot} + \widehat{\mathbf{H}}_c^{(1)\top}\boldsymbol{C}_{j\cdot}) - \boldsymbol{R}_{i\cdot}^\top\mathbf{F}_t\boldsymbol{C}_{j\cdot} \\ &= (\widehat{\boldsymbol{R}}_{i\cdot}^{(1)} - \widehat{\mathbf{H}}_r^{(1)\top}\boldsymbol{R}_{i\cdot})^\top\widehat{\mathbf{F}}_t^{(2)}(\widehat{\boldsymbol{C}}_{j\cdot}^{(1)} - \widehat{\mathbf{H}}_c^{(1)\top}\boldsymbol{C}_{j\cdot}) + \boldsymbol{R}_{i\cdot}^\top\widehat{\mathbf{H}}_r^{(1)}\widehat{\mathbf{F}}_t^{(2)}(\widehat{\boldsymbol{C}}_{j\cdot}^{(1)} - \widehat{\mathbf{H}}_c^{(1)\top}\boldsymbol{C}_{j\cdot}) \\ &+ (\widehat{\boldsymbol{R}}_{i\cdot}^{(1)} - \widehat{\mathbf{H}}_r^{(1)\top}\boldsymbol{R}_{i\cdot})^\top\widehat{\mathbf{F}}_t^{(2)}\widehat{\mathbf{H}}_c^{(1)\top}\boldsymbol{C}_{j\cdot} + \boldsymbol{R}_{i\cdot}^\top(\widehat{\mathbf{H}}_r^{(1)}\widehat{\mathbf{F}}_t^{(2)}\widehat{\mathbf{H}}_c^{(1)\top} - \mathbf{F}_t)\boldsymbol{C}_{j\cdot}.\end{aligned}$$

然后，根据柯西-施瓦茨不等式、定理 3.1、定理 3.2 以及估计量的一致性

因子矩阵，我们有

$$|\widehat{S}^{(2)}_{t,ij}-S_{t,ij}| = O_p\left(\frac{1}{\sqrt{Tp_1}}+\frac{1}{\sqrt{p_1p_2}}+\frac{1}{\sqrt{Tp_2}\nu_{\min}(\mathbf{H}_2)}+\frac{1}{\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu^2_{\min}(\mathbf{H}_2)}+\frac{1}{p_1p_2\nu_{\min}(\mathbf{H}_1)\nu^2_{\min}(\mathbf{H}_2)}\right).$$

□

S1.6 定理证明 3.4

证明。首先回顾

$$\widehat{\mathbf{R}}^{(s+1)} = \sqrt{p_1}\left(\sum_{t=1}^T \mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{F}}_t^{(s+1)\top}\right)\left[\left(\sum_{t=1}^T \widehat{\mathbf{F}}_t^{(s+1)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{X}_t^\top\right)\left(\sum_{t=1}^T \mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{F}}_t^{(s+1)\top}\right)\right]^{-1/2}.$$

$$\begin{aligned}
\sum_{t=1}^T \mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{F}}_t^{(s+1)\top} &= \frac{1}{p_1p_2}\sum_{t=1}^T \mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s)} \\
&= \frac{1}{p_1p_2}\sum_{t=1}^T (\mathbf{R}\mathbf{F}_t\mathbf{C}^\top+\mathbf{E}_t)\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}(\mathbf{R}\mathbf{F}_t\mathbf{C}^\top+\mathbf{E}_t)^\top\widehat{\mathbf{R}}^{(s)} \\
&= \frac{1}{p_1p_2}\sum_{t=1}^T \mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s)} + \frac{1}{p_1p_2}\sum_{t=1}^T \mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)} \\
&+ \frac{1}{p_1p_2}\sum_{t=1}^T \mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s)} + \frac{1}{p_1p_2}\sum_{t=1}^T \mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)} \\
&= \boldsymbol{\delta}_1^{(s+1)}+\boldsymbol{\delta}_2^{(s+1)}+\boldsymbol{\delta}_3^{(s+1)}+\boldsymbol{\delta}_4^{(s+1)}.
\end{aligned}$$

As

$$\begin{aligned}
\|\frac{1}{T}\boldsymbol{\delta}_1^{(s+1)}\|_F^2 &= \|\frac{1}{Tp_1p_2}\sum_{t=1}^T \mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s)}\|_F^2 \\
&\asymp p_2^2\|\frac{1}{T}\sum_{t=1}^T \mathbf{R}\mathbf{F}_t\mathbf{F}_t^\top\|_F^2 \le p_2^2\|\mathbf{R}\|_F^2\|\frac{1}{T}\sum_{t=1}^T \mathbf{F}_t\mathbf{F}_t^\top\|_F^2 = O_p\left(p_1p_2^2\right),
\end{aligned}$$

我们还可以证明 $\|\boldsymbol{\delta}_1^{(s+1)}\|_F^2 \gtrsim T^2p_1p_2^2$，因此，$\|\boldsymbol{\delta}_1^{(s+1)}\|_F^2 \asymp O_p\left(T^2p_1p_2^2\right)$。

根据引理 S2.9 （1）和（2），可得:

$$\begin{aligned}
\|\boldsymbol{\delta}_2^{(s+1)}\|_F^2 &= \|\frac{1}{p_1p_2}\sum_{t=1}^T \mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\|_F^2 \asymp \frac{1}{p_1^2}\|\sum_{t=1}^T \mathbf{R}\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\|_F^2 \\
&= \frac{1}{p_1^2}\|\sum_{t=1}^T \mathbf{R}\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)})\|_F^2 + \frac{1}{p_1^2}\|\sum_{t=1}^T \mathbf{R}\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2 \\
&\le \frac{1}{p_1^2}\|\mathbf{R}\|_F^2\|\sum_{t=1}^T \mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\|_F^2\|\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2 + \frac{1}{p_1^2}\|\mathbf{R}\|_F^2\|\sum_{t=1}^T \mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\mathbf{R}\|_F^2\|\widehat{\mathbf{H}}_r^{(s)}\|_F^2 \\
&= O_p\left(Tp_1p_2^2w_r^{(s)}w_c^{(s)}+Tp_1p_2w_r^{(s)}+Tp_2^2w_c^{(s)}+Tp_2\right).
\end{aligned}$$

同样,

$$\|\boldsymbol{\delta}_3^{(s+1)}\|_F^2 = \|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s)}\|_F^2 \asymp \|\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\|_F^2 = O_p\left(Tp_1p_2^2w_c^{(s)} + Tp_1p_2\right).$$

$$\begin{aligned}
\|\boldsymbol{\delta}_4^{(s+1)}\|_F^2 &= \|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\|_F^2 \\
&\le \|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)})\|_F^2 + \|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2 \\
&\le \frac{1}{p_1^2p_2^2}\left(\sum_{t=1}^{T}\|\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\|_F^2\right)^2\|\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2 + \frac{1}{p_1^2p_2^2}\|\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\mathbf{R}\|_F^2\|\widehat{\mathbf{H}}_r^{(s)}\|_F^2 \\
&= O_p\left(Tp_2 + \frac{T^2}{p_1} + T^2p_1w_r^{(s)} + T^2p_1p_2w_r^{(s)}w_c^{(s)} + T^2p_1p_2^2w_r^{(s)}w_c^{(s)2} + \frac{T^2p_2^2}{p_1}w_c^{(s)2} + Tp_2^2w_c^{(s)2}\right),
\end{aligned}$$

最后一个等号是根据引理 S2.7 (1) 和引理 S2.8 (1) 得出的。

令 $\mathsf{Z}^{(s+1)} = \left(\sum_{t=1}^{T}\mathsf{F}_t^{(s+1)}\mathsf{C}^{(s)\top}\mathsf{X}_t^\top\right)\left(\sum_{t=1}^{T}\mathsf{X}_t\mathsf{C}^{(s)}\mathsf{F}_t^{(s+1)\top}\right)$,

$$\begin{aligned}
\mathbf{Z}^{(s+1)} &= \frac{1}{p_1^2p_2^2}\left(\sum_{t=1}^{T}\widehat{\mathbf{R}}^{(s)\top}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{X}_t^\top\right)\left(\sum_{t=1}^{T}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s)}\right) \\
&= \sum_{i=1}^{4}\sum_{j=1}^{4}\boldsymbol{\delta}_i^{(s+1)\top}\boldsymbol{\delta}_j^{(s+1)},
\end{aligned}$$

于是我们可以证明 $\|(\mathbf{Z}^{(s+1)})^{-1/2}\|_F^2 = O_p\left(\frac{1}{T^2p_1p_2^2}\right)$.

$$\begin{aligned}
\widehat{\mathbf{R}}^{(s+1)} &= \sqrt{p_1}\left(\sum_{t=1}^{T}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{F}}_t^{(s+1)\top}\right)\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^{(s+1)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{X}_t^\top\right)\left(\sum_{t=1}^{T}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{F}}_t^{(s+1)\top}\right)\right]^{-1/2} \\
&= \sqrt{p_1}\left(\sum_{t=1}^{T}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{F}}_t^{(s+1)\top}\right)\left(\mathbf{Z}^{(s+1)}\right)^{1/2} \\
&= \mathbf{R}\left[\frac{1}{\sqrt{p_1}p_2}\left(\sum_{t=1}^{T}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s)}\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\right] \\
&+ \frac{1}{\sqrt{p_1}p_2}\left(\sum_{t=1}^{T}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2} \\
&= \frac{1}{\sqrt{p_1}p_2}\left(\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s)}\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2} \\
&+ \frac{1}{\sqrt{p_1}p_2}\left(\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}.
\end{aligned}$$

记 $\widehat{\mathrm{H}}_r^{(s+1)} = \frac{1}{\sqrt{p_1}p_2}\left(\sum_{t=1}^T \mathrm{F}_t\mathrm{C}^\top \widehat{\mathrm{C}}^{(s)} \widehat{\mathrm{C}}^{(s)\top}\mathrm{C}\mathrm{F}_t^\top \mathrm{R}^\top \widehat{\mathrm{R}}^{(s)}\right)\left(\mathrm{Z}^{(s+1)}\right)^{-1/2}$，则有

$$\begin{aligned}
\widehat{\mathbf{R}}^{(s+1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)} &= \frac{1}{\sqrt{p_1}p_2}\left(\sum_{t=1}^T \mathbf{R}\mathbf{F}_t\mathbf{C}^\top \widehat{\mathbf{C}}^{(s)} \widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top \widehat{\mathbf{R}}^{(s)}\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2} \\
&+ \frac{1}{\sqrt{p_1}p_2}\left(\sum_{t=1}^T \mathbf{E}_t\widehat{\mathbf{C}}^{(s)} \widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top \mathbf{R}^\top \widehat{\mathbf{R}}^{(s)}\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2} \\
&+ \frac{1}{\sqrt{p_1}p_2}\left(\sum_{t=1}^T \mathbf{E}_t\widehat{\mathbf{C}}^{(s)} \widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top \widehat{\mathbf{R}}^{(s)}\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2} \\
&= \sqrt{p_1}(\boldsymbol{\delta}_2^{(s+1)} + \boldsymbol{\delta}_3^{(s+1)} + \boldsymbol{\delta}_4^{(s+1)})(\mathbf{Z}^{(s+1)})^{-1/2},
\end{aligned}$$

因此，

$$\frac{1}{p_1}\|\widehat{\mathbf{R}}^{(s+1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}\|_F^2 = O_p\left(\frac{1}{Tp_2} + \frac{1}{p_1^2p_2^2} + \frac{1}{p_2^2}w_r^{(s)} + \frac{1}{T}w_c^{(s)} + \frac{w_r^{(s)}w_c^{(s)}}{p_2} + w_r^{(s)}w_c^{(s)2} + \frac{1}{p_1^2}w_c^{(s)2}\right).$$

对于 $\widehat{\mathrm{R}}^{(s+1)}$ 的行一致性，我们有

$$\begin{aligned}
\widehat{\boldsymbol{R}}_{i\cdot}^{(s+1)} - \widehat{\mathbf{H}}_r^{(s+1)\top}\boldsymbol{R}_{i\cdot} &= \frac{1}{\sqrt{p_1}p_2}\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\left(\sum_{t=1}^T \widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)} \widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top \boldsymbol{R}_{i\cdot}\right) \\
&+ \frac{1}{\sqrt{p_1}p_2}\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\left(\sum_{t=1}^T \widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)} \widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right) \\
&+ \frac{1}{\sqrt{p_1}p_2}\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\left(\sum_{t=1}^T \widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)} \widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right).
\end{aligned}$$

首先，由引理 S2.9 (1) 和 (2),

$$\begin{aligned}
\left\|\sum_{t=1}^T \widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)} \widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top \boldsymbol{R}_{i\cdot}\right\|_F^2 &\asymp p_2^2\left\|\sum_{t=1}^T \widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\right\|_F^2 \\
&\le p_2^2\left\|\sum_{t=1}^T \mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\right\|_F^2\left\|\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\right\|_F^2 + p_2^2\left\|\sum_{t=1}^T \mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\right\|_F^2 \\
&= O_p\left(Tp_1^2p_2^4w_r^{(s)}w_c^{(s)} + Tp_1^2p_2^3w_r^{(s)} + Tp_1p_2^4w_c^{(s)} + Tp_1p_2^3\right),
\end{aligned}$$

那么，

$$\begin{aligned}
\left\|\frac{1}{\sqrt{p_1}p_2}\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\left(\sum_{t=1}^T \widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)} \widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top \boldsymbol{R}_{i\cdot}\right)\right\|_F^2 &\le \frac{1}{p_1p_2^2}\left\|(\mathbf{Z}^{(s+1)})^{-1/2}\right\|_F^2\left\|\sum_{t=1}^T \widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)} \widehat{\mathbf{C}}^{(s)\top}\mathbf{C}\mathbf{F}_t^\top \boldsymbol{R}_{i\cdot}\right\|_F^2 \\
&= O_p\left(\frac{w_r^{(s)}w_c^{(s)}}{T} + \frac{w_r^{(s)}}{Tp_2} + \frac{w_c^{(s)}}{Tp_1} + \frac{1}{Tp_1p_2}\right).
\end{aligned}$$

其次，与引理的证明类似S2.9 (1)，我们有

$$\left\|\sum_{t=1}^{T}\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2=O_p\left(Tp_2^2w_c^{(s)}+Tp_2\right),$$

$$\begin{aligned}\left\|\frac{1}{\sqrt{p_1}p_2}\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\left(\sum_{t=1}^{T}\widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right)\right\|_F^2&\asymp p_1\left\|\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\sum_{t=1}^{T}\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2\\&\le p_1\left\|\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\right\|_F^2\left\|\sum_{t=1}^{T}\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2\\&=O_p\left(\frac{w_c^{(s)}}{T}+\frac{1}{Tp_2}\right).\end{aligned}$$

第三，与引理 S2.7 (1) 和引理 S2.8 (1) 的证明类似，我们有

$$\left\|\sum_{t=1}^{T}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2=O_p\left(Tp_2^2w_c^{(s)}+Tp_2\right),$$

$$\left\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2=O_p\left(Tp_1p_2^3+T^2p_2^2+(T^2p_2^4+Tp_1p_2^4)w_c^{(s)2}\right),$$

因此，

$$\begin{aligned}&\left\|\sum_{t=1}^{T}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2\le\left\|\sum_{t=1}^{T}(\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)})^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2+\left\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2\\&\le\left\|\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\right\|_F^2\left\|\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\right\|_F^2\left\|\sum_{t=1}^{T}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2+\left\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2\\&=O_p\left(Tp_1p_2^3+T^2p_2^2+T^2p_1^2p_2^2w_r^{(s)}+Tp_1p_2^4w_c^{(s)2}+T^2p_2^4w_c^{(s)2}+T^2p_1^2p_2^4w_c^{(s)2}w_r^s+T^2p_1^2p_2^3w_r^{(s)}w_c^{(s)}\right)\end{aligned}$$

$$\begin{aligned}&\left\|\frac{1}{\sqrt{p_1}p_2}\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\left(\sum_{t=1}^{T}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right)\right\|_F^2\le\frac{1}{p_1p_2^2}\left\|\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\right\|_F^2\left\|\sum_{t=1}^{T}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\boldsymbol{e}_{t,i\cdot}\right\|_F^2\\&=O_p\left(\frac{1}{Tp_1p_2}+\frac{1}{p_1^2p_2^2}+\frac{w_r^{(s)}}{p_2^2}+\frac{w_c^{(s)2}}{Tp_1}+\frac{w_c^{(s)2}}{p_1^2}+w_r^{(s)}w_c^{(s)2}+\frac{w_r^{(s)}w_c^{(s)}}{p_2}\right).\end{aligned}$$

因此

$$\|\widehat{\boldsymbol{R}}_{i\cdot}^{(s+1)}-\widehat{\mathbf{H}}_r^{(s+1)\top}\boldsymbol{R}_{i\cdot}\|_2^2=O_p\left(\frac{1}{Tp_2}+\frac{1}{p_1^2p_2^2}+\frac{1}{p_2^2}w_r^{(s)}+\frac{1}{T}w_c^{(s)}+\frac{w_r^{(s)}w_c^{(s)}}{p_2}+w_r^{(s)}w_c^{(s)2}+\frac{1}{p_1^2}w_c^{(s)2}\right).$$

至于 $\widehat{\mathbf{C}}^{(s+1)}$,

$$\widehat{\mathbf{C}}^{(s+1)} = \sqrt{p_2}\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{F}}_t^{(s+1)}\right)\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^{(s+1)\top}\widehat{\mathbf{R}}^{(s+1)\top}\mathbf{X}_t\right)\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{F}}_t^{(s+1)}\right)\right]^{-1/2}.$$

$$\begin{aligned}
\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{F}}_t^{(s+1)} &= \frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)}\\
&= \frac{1}{p_1p_2}\sum_{t=1}^{T}(\mathbf{R}\mathbf{F}_t\mathbf{C}^\top+\mathbf{E}_t)^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}(\mathbf{R}\mathbf{F}_t\mathbf{C}^\top+\mathbf{E}_t)\widehat{\mathbf{C}}^{(s)}\\
&= \frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)} + \frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\\
&+ \frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)} + \frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\\
&= \boldsymbol{\Delta}_1^{(s+1)}+\boldsymbol{\Delta}_2^{(s+1)}+\boldsymbol{\Delta}_3^{(s+1)}+\boldsymbol{\Delta}_4^{(s+1)}.
\end{aligned}$$

As

$$\begin{aligned}
\|\frac{1}{T}\boldsymbol{\Delta}_1^{(s+1)}\|_F &= \|\frac{1}{Tp_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\|_F^2\\
&\asymp p_1^2\|\frac{1}{T}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{F}_t\|_F^2 = O_p\left(p_1^2p_2\right),
\end{aligned}$$

$\|\boldsymbol{\Delta}_1^{(s+1)}\|_F^2 \geq \|\boldsymbol{\Delta}_1^{(s+1)}\|_2^2 \gtrsim T^2p_1^2p_2$, 则为 $\|\boldsymbol{\Delta}_1^{(s+1)}\|_F^2 \asymp O_p\left(T^2p_1^2p_2\right)$。

由引理 S2.9 (3),

$$\|\boldsymbol{\Delta}_2^{(s+1)}\|_F^2 = \|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\|_F^2 \asymp \|\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\mathbf{F}_t\|_F = O_p\left(Tp_1^2p_2w_r^{(s+1)}+Tp_1p_2\right).$$

同样地，根据引理 S2.9 （3）和（4），可得

$$\begin{aligned}
\|\boldsymbol{\Delta}_3^{(s+1)}\|_F^2 &= \|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\|_F^2 \asymp \frac{1}{p_2^2}\|\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\|_F^2\\
&\leq \frac{1}{p_2^2}\|\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})\|_F^2 + \frac{1}{p_2^2}\|\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2\\
&\leq \frac{1}{p_2^2}\|\mathbf{C}\|_F^2\|\sum_{t=1}^{T}\mathbf{F}_t^\top\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\|_F^2\|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2 + \frac{1}{p_2^2}\|\mathbf{C}\|_F^2\|\sum_{t=1}^{T}\mathbf{F}_t^\top\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\mathbf{C}\|_F^2\|\widehat{\mathbf{H}}_c^{(s)}\|_F^2\\
&= O_p\left(Tp_1^2p_2w_r^{(s)}w_c^{(s)}+Tp_1p_2w_c^{(s)}+Tp_1^2w_r^{(s)}+Tp_1\right).
\end{aligned}$$

根据引理 S2.7 (2) 和引理 S2.8 (2)，我们有

$$
\begin{aligned}
\|\mathbf{\Delta}_4^{(s+1)}\|_F^2 &= \|\frac{1}{p_1p_2}\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\|_F^2 \\
&\le \frac{1}{p_1^2p_2^2}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})\|_F^2+\frac{1}{p_1^2p_2^2}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2 \\
&\le \frac{1}{p_1^2p_2^2}\sum_{t=1}^{T}\|\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\|_F^2\sum_{t=1}^{T}\|\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\|_F^2\|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2+\frac{1}{p_1^2p_2^2}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\mathbf{C}\|_F^2\|\widehat{\mathbf{H}}_c^{(s)}\|_F^2 \\
&= O_p\left(Tp_1+\frac{T^2}{p_2}+T^2p_2w_c^{(s)}+T^2p_1p_2w_r^{(s)}w_c^{(s)}+T^2p_1^2p_2w_r^{(s)}w_r^{(s+1)}w_c^{(s)}+Tp_1^2w_r^{(s)}w_r^{(s+1)}+\frac{T^2p_1^2}{p_2}w_r^{(s)}w_r^{(s+1)}\right).
\end{aligned}
$$

令 $\mathrm{Y}^{(s+1)}=\left(\sum_{t=1}^{T}\widehat{\mathrm{F}}_t^{(s+1)\top}\widehat{\mathrm{R}}^{(s+1)\top}\mathrm{X}_t\right)\left(\sum_{t=1}^{T}\mathrm{X}_t^\top\widehat{\mathrm{R}}^{(s+1)}\widehat{\mathrm{F}}_t^{(s+1)}\right)$，则 $\mathrm{Y}^{(s+1)}=\sum_{i=1}^{4}\sum_{j=1}^{4}\mathbf{\Delta}_i^{(s+1)\top}\mathbf{\Delta}_j^{(s+1)}$，我们可以证明

$$
\left\|\left(\mathbf{Y}^{(s+1)}\right)^{-1/2}\right\|_F=O_p\left(\frac{1}{T^2p_1^2p_2}\right).
$$

$$
\begin{aligned}
\widehat{\mathbf{C}}^{(s+1)} &= \sqrt{p_2}\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{F}}_t^{(s+1)}\right)\left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^{(s+1)\top}\widehat{\mathbf{R}}^{(s+1)\top}\mathbf{X}_t\right)\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{F}}_t^{(s+1)}\right)\right]^{-1/2} \\
&= \sqrt{p_2}\left(\sum_{t=1}^{T}\mathbf{X}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{F}}_t^{(s+1)}\right)\left(\mathbf{Y}^{(s+1)}\right)^{-1/2} \\
&= \mathbf{C}\left[\frac{1}{p_1\sqrt{p_2}}\left(\sum_{t=1}^{T}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\right)\left(\mathbf{Y}^{(s+1)}\right)^{-1/2}\right] \\
&+ \frac{1}{p_1\sqrt{p_2}}\left(\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\right)\left(\mathbf{Y}^{(s+1)}\right)^{-1/2} \\
&+ \frac{1}{p_1\sqrt{p_2}}\left(\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\right)\left(\mathbf{Y}^{(s+1)}\right)^{-1/2} \\
&+ \frac{1}{p_1\sqrt{p_2}}\left(\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\right)\left(\mathbf{Y}^{(s+1)}\right)^{-1/2}.
\end{aligned}
$$

记 $\widehat{\mathrm{H}}_c^{(s+1)}=\frac{1}{p_1\sqrt{p_2}}\left(\sum_{t=1}^{T}\mathrm{F}_t^\top\mathrm{R}^\top\widehat{\mathrm{R}}^{(s+1)}\widehat{\mathrm{R}}^{(s)\top}\mathrm{R}\mathrm{F}_t\mathrm{C}^\top\widehat{\mathrm{C}}^{(s)}\right)\left(\mathrm{Y}^{(s+1)}\right)^{-1/2}$，则

$$
\begin{aligned}
\widehat{\mathbf{C}}^{(s+1)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s+1)} &= \frac{1}{p_1\sqrt{p_2}}\left(\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)}\right)\left(\mathbf{Y}^{(s+1)}\right)^{-1/2} \\
&+ \frac{1}{p_1\sqrt{p_2}}\left(\sum_{t=1}^{T}\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\right)\left(\mathbf{Y}^{(s+1)}\right)^{-1/2} \\
&+ \frac{1}{p_1\sqrt{p_2}}\left(\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\right)\left(\mathbf{Y}^{(s+1)}\right)^{-1/2} \\
&= \sqrt{p_2}(\mathbf{\Delta}_2^{(s+1)}+\mathbf{\Delta}_3^{(s+1)}+\mathbf{\Delta}_4^{(s+1)})(\mathbf{Y}^{(s+1)})^{-1/2}.
\end{aligned}
$$

thus

$$\frac{1}{p_2}\|\widehat{\mathbf{C}}^{(s+1)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(s+1)}\|_F^2 = O_p\left(\frac{1}{Tp_1} + \frac{1}{p_1^2p_2^2} + \gamma_c^{(s+1)}\right),$$

其中 $\gamma_c^{(s+1)} = \frac{1}{Tp_2}w_r^{(s)} + \frac{1}{T}w_r^{(s+1)} + \frac{1}{p_1^2}w_c^{(s)} + \frac{1}{p_1}w_r^{(s)}w_c^{(s)} + \frac{1}{T}w_r^{(s)}w_c^{(s)} + w_r^{(s)}w_r^{(s+1)}w_c^{(s)} + \frac{1}{p_2^2}w_r^{(s)}w_r^{(s+1)}$。

与证明 $\widehat{\mathbf{R}}^{(s+1)}$ 的行一致性类似，对于任意 $j \in [p_2]$，我们可以得到

$$\|\widehat{\boldsymbol{C}}_{j\cdot}^{(s+1)} - \widehat{\mathbf{H}}_c^{(s+1)\top}\boldsymbol{C}_{j\cdot}\|_2^2 = O_p\left(\frac{1}{Tp_1} + \frac{1}{p_1^2p_2^2} + \gamma_c^{(s+1)}\right),$$

where $\gamma_c^{(s+1)} = \frac{1}{Tp_2}w_r^{(s)} + \frac{1}{T}w_r^{(s+1)} + \frac{1}{p_1^2}w_c^{(s)} + \frac{1}{p_1}w_r^{(s)}w_c^{(s)} + \frac{1}{T}w_r^{(s)}w_c^{(s)} + w_r^{(s)}w_r^{(s+1)}w_c^{(s)} + \frac{1}{p_2^2}w_r^{(s)}w_r^{(s+1)}$.

□

S1.7 定理证明 3.5

证明。根据定义，

$$\begin{aligned}
\widehat{\mathbf{F}}_t^{(s+1)} &= \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(s)\top}\mathbf{X}_t\widehat{\mathbf{C}}^{(s)} = \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(s)\top}\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\widehat{\mathbf{C}}^{(s)} + \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)} \\
&= \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(s)\top}\left(\mathbf{R} - \widehat{\mathbf{R}}^{(s)}(\widehat{\mathbf{H}}_r^{(s)})^{-1} + \widehat{\mathbf{R}}^{(s)}(\widehat{\mathbf{H}}_r^{(s)})^{-1}\right)\mathbf{F}_t\left(\mathbf{C} - \widehat{\mathbf{C}}^{(s)}(\widehat{\mathbf{H}}_c^{(s)})^{-1} + \widehat{\mathbf{C}}^{(s)}(\widehat{\mathbf{H}}_c^{(s)})^{-1}\right)^\top\widehat{\mathbf{C}}^{(s)} \\
&+ \frac{1}{p_1p_2}\left(\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)} + \mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\right)^\top\mathbf{E}_t\left(\widehat{\mathbf{C}}^{(s)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(s)} + \mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\right).
\end{aligned}$$

注意，$\widehat{\mathbf{R}}^{(s)\top}\widehat{\mathbf{R}}^{(s)} = p_1\mathbf{I}_{k_1}$，$\widehat{\mathbf{C}}^{(s)\top}\widehat{\mathbf{C}}^{(s)} = p_2\mathbf{I}_{k_2}$，则

$$\begin{aligned}
\widehat{\mathbf{F}}_t^{(s+1)} - (\widehat{\mathbf{H}}_r^{(s)})^{-1}\mathbf{F}_t\left((\widehat{\mathbf{H}}_c^{(s)})^{-1}\right)^\top &= \frac{1}{p_1p_2}\widehat{\mathbf{R}}^{(s)\top}\left(\mathbf{R} - \widehat{\mathbf{R}}^{(s)}(\widehat{\mathbf{H}}_r^{(s)})^{-1}\right)\mathbf{F}_t\left(\mathbf{C} - \widehat{\mathbf{C}}(\widehat{\mathbf{H}}_c^{(s)})^{-1}\right)^\top\widehat{\mathbf{C}}^{(s)} \\
&+ \frac{1}{p_1}\widehat{\mathbf{R}}^{(s)\top}\left(\mathbf{R} - \widehat{\mathbf{R}}^{(s)}(\widehat{\mathbf{H}}_r^{(s)})^{-1}\right)\mathbf{F}_t\left((\widehat{\mathbf{H}}_c^{(s)})^{-1}\right)^\top \\
&+ \frac{1}{p_2}(\widehat{\mathbf{H}}_r^{(s)})^{-1}\mathbf{F}_t\left(\mathbf{C} - \widehat{\mathbf{C}}(\widehat{\mathbf{H}}_c^{(s)})^{-1}\right)^\top\widehat{\mathbf{C}}^{(s)} \\
&+ \frac{1}{p_1p_2}\left(\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\right)^\top\mathbf{E}_t\left(\widehat{\mathbf{C}}^{(s)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\right) \\
&+ \frac{1}{p_1p_2}\left(\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\right)^\top\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)} + \frac{1}{p_1p_2}\widehat{\mathbf{H}}_r^{(s)\top}\mathbf{R}^\top\mathbf{E}_t\left(\widehat{\mathbf{C}}^{(s)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\right) \\
&+ \frac{1}{p_1p_2}\widehat{\mathbf{H}}_r^{(s)\top}\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}.
\end{aligned}$$

因此，根据柯西-施瓦茨不等式、引理 S2.6 和定理 3.4，我们有

$$\left\|\widehat{\mathbf{F}}_t^{(s+1)} - (\widehat{\mathbf{H}}_r^{(s)})^{-1}\mathbf{F}_t\left((\widehat{\mathbf{H}}_c^{(s)})^{-1}\right)^\top\right\|_F = O_p\left(\frac{\sqrt{w_r^{(s)}}}{\sqrt{p_2}} + \frac{\sqrt{w_r^{(s-1)}}}{\sqrt{Tp_2}} + \frac{\sqrt{w_c^{(s)}}}{\sqrt{p_1}} + \frac{\sqrt{w_c^{(s-1)}}}{\sqrt{Tp_1}} + \frac{1}{\sqrt{p_1p_2}} + \gamma_f^{(s+1)}\right),$$

where $\gamma_f^{(s+1)} = \sqrt{w_r^{(s)}w_c^{(s)}} + \frac{\sqrt{w_r^{(s-1)}w_c^{(s-1)}}}{\sqrt{T}} + \frac{\sqrt{w_r^{(s-1)}}w_c^{(s-1)}}{p_1}$.

接下来，对于任意 t,i,j

$$\begin{aligned}\widehat{S}_{t,ij}^{(s+1)} - S_{t,ij} &= \widehat{\boldsymbol{R}}_{i\cdot}^{(s)\top}\widehat{\mathbf{F}}_t^{(s+1)}\widehat{\boldsymbol{C}}_{j\cdot}^{(s)} - \boldsymbol{R}_{i\cdot}^{\top}\mathbf{F}_t\boldsymbol{C}_{j\cdot} \\ &= (\widehat{\boldsymbol{R}}_{i\cdot}^{(s)} - \widehat{\mathbf{H}}_r^{(s)\top}\boldsymbol{R}_{i\cdot} + \widehat{\mathbf{H}}_r^{(s)\top}\boldsymbol{R}_{i\cdot})^{\top}\widehat{\mathbf{F}}_t^{(s+1)}(\widehat{\boldsymbol{C}}_{j\cdot}^{(s)} - \widehat{\mathbf{H}}_c^{(s)\top}\boldsymbol{C}_{j\cdot} + \widehat{\mathbf{H}}_c^{(s)\top}\boldsymbol{C}_{j\cdot}) - \boldsymbol{R}_{i\cdot}^{\top}\mathbf{F}_t\boldsymbol{C}_{j\cdot} \\ &= (\widehat{\boldsymbol{R}}_{i\cdot}^{(s)} - \widehat{\mathbf{H}}_r^{(s)\top}\boldsymbol{R}_{i\cdot})^{\top}\widehat{\mathbf{F}}_t^{(s+1)}(\widehat{\boldsymbol{C}}_{j\cdot}^{(s)} - \widehat{\mathbf{H}}_c^{(s)\top}\boldsymbol{C}_{j\cdot}) + \boldsymbol{R}_{i\cdot}^{\top}\widehat{\mathbf{H}}_r^{(s)}\widehat{\mathbf{F}}_t^{(s+1)}(\widehat{\boldsymbol{C}}_{j\cdot}^{(s)} - \widehat{\mathbf{H}}_c^{(s)\top}\boldsymbol{C}_{j\cdot}) \\ &+ (\widehat{\boldsymbol{R}}_{i\cdot}^{(s)} - \widehat{\mathbf{H}}_r^{(s)\top}\boldsymbol{R}_{i\cdot})^{\top}\widehat{\mathbf{F}}_t^{(s+1)}\widehat{\mathbf{H}}_c^{(s)\top}\boldsymbol{C}_{j\cdot} + \boldsymbol{R}_{i\cdot}^{\top}(\widehat{\mathbf{H}}_r^{(s)}\widehat{\mathbf{F}}_t^{(s+1)}\widehat{\mathbf{H}}_c^{(s)\top} - \mathbf{F}_t)\boldsymbol{C}_{j\cdot}.\end{aligned}$$

然后，由柯西-施瓦茨不等式、定理 3.1、定理 3.2 以及因子矩阵估计量的一致性，可得

$$|\widehat{S}_{t,ij}^{(s+1)} - S_{t,ij}| = O_p\left(\sqrt{w_r^{(s)}} + \frac{\sqrt{w_r^{(s-1)}}}{\sqrt{Tp_2}} + \sqrt{w_c^{(s)}} + \frac{\sqrt{w_c^{(s-1)}}}{\sqrt{Tp_1}} + \frac{1}{\sqrt{p_1p_2}} + \gamma^{(s+1)}\right),$$

where $\gamma^{(s+1)} = \dfrac{\sqrt{w_r^{(s-1)}w_c^{(s-1)}}}{\sqrt{T}} + \dfrac{\sqrt{w_r^{(s-1)}}w_c^{(s-1)}}{p_1}$.

□

S1.8 定理 3.6 的证明

证明。 类似于定理 2.1 的证明，对于任意有界的 $m_1 \geq k_1$， $m_2 \geq k_2$，有

$$\|\widetilde{\mathbf{F}}_t - \widetilde{\mathbf{H}}_1\mathbf{F}_t\widetilde{\mathbf{H}}_2^{\top}\|_F = \|\widetilde{\boldsymbol{\mathcal{E}}}_t\|_F = O_p\left(\frac{1}{\sqrt{p_1}}\right),$$

其中 $\widetilde{\mathrm{H}}_1 = \widehat{\mathrm{R}}^{(S-1)\top}\mathrm{R}/p_1$, $\widetilde{\mathrm{H}}_2 = \widehat{\mathrm{C}}^{(S-1)\top}\mathrm{C}/p_2$, $\widetilde{\mathrm{E}}_t = \widehat{\mathrm{R}}^{(S-1)\top}\mathrm{E}_t\widehat{\mathrm{C}}^{(S-1)}/(p_1p_2)$，$S$ 是直到收敛或达到预设的最大迭代次数为止的迭代次数。因此

$$\left\|\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^{\top} - \frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{H}}_1\mathbf{F}_t\widetilde{\mathbf{H}}_2^{\top}\widetilde{\mathbf{H}}_2\mathbf{F}_t^{\top}\widetilde{\mathbf{H}}_1^{\top}\right\|_F = O_p\left(\frac{1}{\sqrt{p_1}}\right).$$

$\widetilde{}$ $\widetilde{}$我们首先证明 $\mathrm{rank}(\mathrm{H}_1) = k_1$、$\mathrm{rank}(\mathrm{H}_2) = k_2$。

记 $\widehat{\mathrm{R}}_{k_1}^{(s)}$ 为 $\widehat{\mathrm{R}}^{(s)}$的前 k_1 列， $\widehat{\mathrm{C}}_{k_2}^{(s)}$ 的定义类似。于是

$$\widehat{\mathbf{R}}_{k_1}^{(s)} = \sqrt{p_1}\left(\sum_{t=1}^{T}\mathbf{X}_t\widehat{\mathbf{C}}^{(s-1)}\widehat{\mathbf{F}}_t^{(s)\top}\right)\mathbf{Z}_{k_1}^{(s)},$$

其中，$\mathrm{Z}_{k_1}^{(s)}$ 是由 k_1 的前 k_1 列组成的矩阵

$$\left(\mathbf{Z}^{(s)}\right)^{-1/2} = \left[\left(\sum_{t=1}^{T}\widehat{\mathbf{F}}_t^{(s)}\widehat{\mathbf{C}}^{(s-1)\top}\mathbf{X}_t^{\top}\right)\left(\sum_{t=1}^{T}\mathbf{X}_t\widehat{\mathbf{C}}^{(s-1)}\widehat{\mathbf{F}}_t^{(s)\top}\right)\right]^{-1/2},$$

因此， $\|\mathrm{Z}_{k_1}^{(s)}\|_2 \leq 1/\lambda_{\min}\left((\mathbf{Z}^{(s)})^{1/2}\right)$。

假设 $k_1 = k_2 = 1$、$m_1 > k_1$、$m_2 > k_2$，以 $s = 1$ 为例。类似于定理的证明

3.1，我们有

$$\nu_{\min}(\boldsymbol{\delta}_1) = \nu_{\min}\left(\frac{\sqrt{p_1}p_2}{T}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{H}_2^\top\mathbf{H}_2\mathbf{F}_t^\top\mathbf{H}_1\right) \gtrsim \sqrt{p_1}p_2\lambda_{\min}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{H}_2^\top\mathbf{H}_2\mathbf{F}_t^\top\right)\nu_{\min}(\mathbf{H}_1)$$

$$\gtrsim \sqrt{p_1}p_2\lambda_{\min}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{F}_t^\top\right)\nu_{\min}(\mathbf{H}_2^\top\mathbf{H}_2)\nu_{\min}(\mathbf{H}_1) \gtrsim \sqrt{p_1}p_2\nu_{\min}^2(\mathbf{H}_2)\nu_{\min}(\mathbf{H}_1).$$

$\|\boldsymbol{\delta}_2\|_F = O_p\left(\sqrt{p_2/T}\nu_{\min}(\mathbf{H}_2)\right), \|\boldsymbol{\delta}_3\|_F = O_p\left(\sqrt{p_1p_2/T}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)\right),$
$\|\boldsymbol{\delta}_4\|_F = O_p\left(\sqrt{p_2/T} + 1/\sqrt{p_1}\right)$，因此

$$\lambda_{\min}\left((\mathbf{Z}^{(s)})^{1/2}\right) \gtrsim \sqrt{p_1}p_2\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2).$$

Let

$$\widehat{\mathbf{H}}_{r,k_1}^{(1)} = \left(\frac{1}{T\sqrt{p_1}p_2}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1\right)\mathbf{Z}_{k_1}^{(1)},$$

因此，我们可以证明

$$\frac{1}{p_1}\|\widehat{\mathbf{R}}_{k_1}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_{r,k_1}^{(1)}\|_F^2 = o_p(1).$$

最后，我们有 $\mathrm{I}_{k_1} = p_1^{-1}\widehat{\mathrm{R}}_{k_1}^{(1)\top}\mathrm{R}\,\widehat{\mathrm{H}}_r^{(1),\,k_1} + o_p(1) = \widehat{\mathrm{H}}_r^{(1)\top,\,k_1}\widehat{\mathrm{H}}_r^{(1),\,k_1} + o_p(1)$。也就是说，$k_1 = \mathrm{rank}(p_1^{-1}\widehat{\mathrm{R}}_{k_1}^{(1)\top}\mathrm{R}) \le \mathrm{rank}(p_1^{-1}\widehat{\mathrm{R}}^{(1)\top}\mathrm{R}) \le k_1$。因此，$\mathrm{rank}(p_1^{-1}\widehat{\mathrm{R}}^{(1)\top}\mathrm{R}) = k_1$。类似地，$\mathrm{rank}(p_2^{-1}\widehat{\mathrm{C}}^{(1)\top}\mathrm{C}) = k_2$。按照类似的$\widetilde{\quad}$证明，我们断言 $\mathrm{rank}(\mathrm{H}_1) = k_1, \mathrm{rank}(\mathrm{H}_2) = k_2$。

由于秩$(\widetilde{\mathrm{H}}_1) = k_1$，秩$(\widetilde{\mathrm{H}}_2) = k_2$，因此，秩$(\widetilde{\mathbf{H}}_1^\top\widetilde{\mathrm{H}}_1) = k_1$，秩$(\widetilde{\mathbf{H}}_2^\top\widetilde{\mathrm{H}}_2) = k_2$，

$$\mathrm{rank}\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{H}}_1\mathbf{F}_t\widetilde{\mathbf{H}}_2^\top\widetilde{\mathbf{H}}_2\mathbf{F}_t^\top\widetilde{\mathbf{H}}_1^\top\right) = \mathrm{rank}\left(\frac{1}{T}\sum_{t=1}^{T}\mathbf{F}_t\mathbf{F}_t^\top\right) = k_1,$$

因此，$T^{-1}\sum_{t=1}^{T}\widetilde{\mathrm{H}}_1\mathrm{F}_t\widetilde{\mathrm{H}}_2^\top\widetilde{\mathrm{H}}_2\mathrm{F}_t^\top\widetilde{\mathrm{H}}_1^\top$恰有 $m_1 - k_1$ 个 $\lambda = 0$的重根。

对于 $i = 1, \ldots, k_1$，由韦尔定理可得

$$\left|\lambda_i\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t\right) - \lambda_i\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{H}}_1\mathbf{F}_t\widetilde{\mathbf{H}}_2^\top\widetilde{\mathbf{H}}_2\mathbf{F}_t^\top\widetilde{\mathbf{H}}_1^\top\right)\right| \le \left\|\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top - \frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{H}}_1\mathbf{F}_t\widetilde{\mathbf{H}}_2^\top\widetilde{\mathbf{H}}_2\mathbf{F}_t^\top\widetilde{\mathbf{H}}_1^\top\right\|_F$$
$$= O_p\left(\frac{1}{\sqrt{p_1}}\right).$$

对于 $i > k_1$，我们有

$$\left|\lambda_i\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t\right)\right| = O_p\left(\frac{1}{\sqrt{p_1}}\right).$$

因此，

$$\lambda_i\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top\right) = \begin{cases}\lambda_i\left(\frac{1}{T}\sum_{t=1}^{T}\widetilde{\mathbf{H}}_1\mathbf{F}_t\widetilde{\mathbf{H}}_2^\top\widetilde{\mathbf{H}}_2\mathbf{F}_t^\top\widetilde{\mathbf{H}}_1^\top\right) + O_p\left(\frac{1}{\sqrt{p_1}}\right), & i \le k_1,\\ O_p\left(\frac{1}{\sqrt{p_1}}\right), & i > k_1.\end{cases}$$

注意，$\|T^{-1}\sum_{t=1}^{T}\widetilde{\mathrm{H}}_1\mathrm{F}_t\widetilde{\mathrm{H}}_2^\top\widetilde{\mathrm{H}}_2\mathrm{F}_t^\top\widetilde{\mathrm{H}}_1^\top\|_F = O_p(1)$ 且 $\mathrm{rank}(T^{-1}\sum_{t=1}^{T}\widetilde{\mathrm{H}}_1\mathrm{F}_t\widetilde{\mathrm{H}}_2^\top\widetilde{\mathrm{H}}_2\mathrm{F}_t^\top\widetilde{\mathrm{H}}_1^\top) = k_1$，由于$T$、$p_1$、$p_2 \to \infty$，我们可以计算特征值比。

$$\max_{i\le k_1-1}\frac{\lambda_i\left(\frac{1}{T}\sum_{t=1}^T\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top\right)}{\lambda_{i+1}\left(\frac{1}{T}\sum_{t=1}^T\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top\right)}=O_p(1),$$

$$\max_{i\ge k_1+1}\frac{\lambda_i\left(\frac{1}{T}\sum_{t=1}^T\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top\right)}{\lambda_{i+1}\left(\frac{1}{T}\sum_{t=1}^T\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top\right)}\le O_p(1),$$

$$\frac{\lambda_{k_1}\left(\frac{1}{T}\sum_{t=1}^T\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top\right)}{\lambda_{k_1+1}\left(\frac{1}{T}\sum_{t=1}^T\widetilde{\mathbf{F}}_t\widetilde{\mathbf{F}}_t^\top\right)}\to\infty,$$

由此可得一致性。

类似地，我们可以得到 $\mathbb{P}(\widehat{k}_2=k_2)\to 1$，因为 $T,p_1,p_2\to\infty$。 □

S2 命题与引理

在本节中，我们首 先给出一些命题和引理，它们对于证明主要定理至关重要 .

命题 S2.1. 假设 $m_1\ge k_1$， $m_2\ge k_2$ 且 T， p_1， $p_2\to\infty$。在假设 2.1-2.3下，我们有

- $\lambda_{\text{最小值}}(\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}})\ge c(p_1p_2)^{-1}$，对于某个 $c>0$，其概率趋于 1。

证明。容易得到

$$\widehat{\mathbf{F}}=\mathbf{F}\mathbf{H}^\top+\boldsymbol{\mathcal{E}},$$

其中 $\mathbf{E}=\mathbf{E}\mathbf{W}/(p_1p_2)$， $\mathbf{W}=\mathbf{W}_2\otimes\mathbf{W}_1$. 令

$$\Delta:=\frac{1}{T}\mathbf{H}\mathbf{F}^\top\boldsymbol{\mathcal{E}}+\frac{1}{T}\boldsymbol{\mathcal{E}}^\top\mathbf{F}\mathbf{H}^\top+\frac{1}{T}\mathbb{E}\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}}+\frac{1}{T}(\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}}-\mathbb{E}\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}}).$$

于是有

$$\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}=\frac{1}{T}\mathbf{H}\mathbf{F}^\top\mathbf{F}\mathbf{H}^\top+\Delta.$$

根据假设

$$\lambda_{\min}\left(\frac{1}{T}\mathbb{E}\mathbf{E}^\top\mathbf{E}\right)=\lambda_{\min}\left(\frac{1}{T}\sum_{t=1}^T\mathbb{E}\mathrm{Vec}(\mathbf{E}_t)\mathrm{Vec}(\mathbf{E}_t)^\top\right)\ge c_0,$$

以及克罗内克积的性质，可以得到

$$\lambda_{\min}\left(\frac{1}{T}\mathbb{E}\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}}\right)\ge\lambda_{\min}\left(\frac{1}{T}\mathbb{E}\mathbf{E}^\top\mathbf{E}\right)\lambda_{\min}\left(\frac{1}{p_1^2p_2^2}\mathbf{W}^\top\mathbf{W}\right)\ge c_0(p_1p_2)^{-1}$$

对某个 $c_0>0$。此外，在以下引理 S2.2 (4) 中，我们证明

$$\|\frac{1}{T}(\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}}-\mathbb{E}\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}})\|_2=O_p\left(\frac{1}{p_1p_2\sqrt{T}}\right).$$

因此，不等式 $\|\frac{1}{T}(\mathbf{E}^\top\mathbf{E} - \mathbb{E}\mathbf{E}^\top\mathbf{E})\|_2 \leq \frac{1}{2}\lambda_{\min}(\frac{1}{T}\mathbb{E}\mathbf{E}^\top\mathbf{E})$ 以高概率成立。现在，我们继续在该事件的条件下进行论证。

令 $\boldsymbol{v}$ 为满足 $\boldsymbol{v}^\top\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}\boldsymbol{v} = \lambda_{\min}(\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}})$ 的单位向量，则 $\boldsymbol{v}^\top\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}\boldsymbol{v} = \boldsymbol{v}^\top\frac{1}{T}\mathbf{H}\mathbf{F}^\top\mathbf{F}\mathbf{H}^\top\boldsymbol{v} + \boldsymbol{v}^\top\Delta\boldsymbol{v}$，因此我们有

$$\lambda_{\min}(\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}) \geq \frac{1}{T}\boldsymbol{v}^\top\mathbf{H}\mathbf{F}^\top\mathbf{F}\mathbf{H}^\top\boldsymbol{v} + \frac{2}{T}\boldsymbol{v}^\top\mathbf{H}\mathbf{F}^\top\boldsymbol{\mathcal{E}}\boldsymbol{v} + \frac{c_0}{2p_1p_2}.$$

若 $\boldsymbol{v}^\top\mathbf{H} = 0$，则 $\lambda_{\min}(\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}) \geq \frac{c_0}{2p_1p_2}$。若 $\boldsymbol{v}^\top\mathbf{H} \neq 0$，由于 $\frac{1}{T}\mathbf{F}^\top\mathbf{F}$ 是正定的，我们以高概率有 $\frac{1}{T}\boldsymbol{v}^\top\mathbf{H}\mathbf{F}^\top\mathbf{F}\mathbf{H}^\top\boldsymbol{v} > 0$。令

$$\alpha_{\boldsymbol{v}}^2 = \frac{1}{T}\boldsymbol{v}^\top\mathbf{H}\mathbf{F}^\top\mathbf{F}\mathbf{H}^\top\boldsymbol{v}, \quad X = (\frac{\alpha_{\boldsymbol{v}}^2}{Tp_1p_2})^{-1/2}2\boldsymbol{v}^\top\frac{1}{T}\mathbf{H}\mathbf{F}^\top\boldsymbol{\mathcal{E}}\boldsymbol{v}, \quad 2\boldsymbol{v}^\top\frac{1}{T}\mathbf{H}\mathbf{F}^\top\boldsymbol{\mathcal{E}}\boldsymbol{v} = X\sqrt{\frac{\alpha_{\boldsymbol{v}}^2}{Tp_1p_2}}.$$

Then

$$\lambda_{\min}(\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}) \geq \alpha_{\boldsymbol{v}}^2 + X\sqrt{\frac{\alpha_{\boldsymbol{v}}^2}{Tp_1p_2}} + \frac{c_0}{2p_1p_2}.$$

首先，我们证明 $X = O_p(1)$。根据假设2.3，有 $\lambda_{\min}(\frac{1}{T}\mathbf{F}^\top\mathbf{F}) > c > 0$，因此我们有

$$\alpha_{\boldsymbol{v}}^2 \geq \lambda_{\min}(\frac{1}{T}\mathbf{F}^\top\mathbf{F})\boldsymbol{v}^\top\mathbf{H}\mathbf{H}^\top\boldsymbol{v} > c\|\boldsymbol{v}^\top\mathbf{H}\|_2^2.$$

根据引理S2.2 (3)，我们有 $\|\frac{1}{T}\mathbf{F}^\top\boldsymbol{\mathcal{E}}\|_2^2 = O_p\left(\frac{1}{Tp_1p_2}\right)$ 因此，

$$|X|^2 \leq 4Tp_1p_2\alpha_{\boldsymbol{v}}^{-2}\|\boldsymbol{v}^\top\mathbf{H}\|_2^2\|\frac{1}{T}\mathbf{F}^\top\boldsymbol{\mathcal{E}}\|_2^2 \leq O_p(1)\alpha_{\boldsymbol{v}}^{-2}\|\boldsymbol{v}^\top\mathbf{H}\|_2^2 \leq O_p(1)c^{-1}\|\boldsymbol{v}^\top\mathbf{H}\|_2^{-2}\|\boldsymbol{v}^\top\mathbf{H}\|_2^2 = O_p(1).$$

接下来我们分两种情况讨论。

情况一：$\alpha_{\boldsymbol{v}}^2 \leq 4|X|\sqrt{\frac{\alpha_{\boldsymbol{v}}^2}{Tp_1p_2}}$. 然后 $|\alpha_{\boldsymbol{v}}| \leq 4|X|\frac{1}{\sqrt{Tp_1p_2}}$和

$$\lambda_{\min}(\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}) \geq \frac{c_0}{2p_1p_2} - |X||\alpha_{\boldsymbol{v}}|\frac{1}{\sqrt{Tp_1p_2}} \geq \frac{c_0}{2p_1p_2} - 4|X|^2\frac{1}{Tp_1p_2} \geq \frac{c_0}{4p_1p_2},$$

其中，最后一个不等式根据 $X = O_p(1)$和 $T \to \infty$这一事实，以趋近于 1 的概率成立。

情况二：$\alpha_{\boldsymbol{v}}^2 > 4|X|\sqrt{\frac{\alpha_{\boldsymbol{v}}^2}{Tp_1p_2}}$，那么

$$\lambda_{\min}(\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}) \geq \alpha_{\boldsymbol{v}}^2 - |X|\sqrt{\frac{\alpha_{\boldsymbol{v}}^2}{Tp_1p_2}} + \frac{c_0}{2p_1p_2} \geq \frac{3}{4}\alpha_{\boldsymbol{v}}^2 + \frac{c_0}{2p_1p_2} \geq \frac{c_0}{2p_1p_2}.$$

根据 如上所述，我们得出结论：概率趋于 1 时，我们有 e

$$\lambda_{\min}(\frac{1}{T}\widehat{\mathbf{F}}^\top\widehat{\mathbf{F}}) > c_0/(p_1p_2).$$

□

引理 S2.2。对于任意 $m_1 \geq 1$ 和 $m_2 \geq 1$，（注意，m_1、m_2 可以小于、等于或大于 k_1、k_2），

(1) $\|\mathbb{E}(\mathsf{E}_t^\top \mathsf{E}_t)\|_2 = O\left(\frac{1}{p_1}\right)$ 且 $\|\mathsf{E}_t\|_2 = O_p\left(\frac{1}{\sqrt{p_1}}\right), t \in [T]$,

(2) $\|\mathsf{E}\|_2 = O_p\left(\sqrt{\frac{T}{p_1 p_2}}\right)$, (3) $\mathbb{E}\|\frac{1}{T}\mathsf{F}^\top \mathsf{E}\|_2^2 \leq O\left(\frac{1}{T p_1 p_2}\right)$,

(4) $\left\|\frac{1}{T}(\mathsf{E}^\top \mathsf{E} - \mathbb{E}\mathsf{E}^\top \mathsf{E})\right\|_2 \leq O_p\left(\frac{1}{p_1 p_2 \sqrt{T}}\right)$.

在此引理中，所有 $\mathbb{E}(\cdot)$ $\mathbb{E}(\cdot|\cdot)$ 和 $\mathsf{Cov}(\cdot)$ 均是在给定 W_1,W_2 的条件下计算的。该引理的证明如下。

证明。 (1) 根据假设 2.3，我们有 $\|\mathbb{E}(\mathbf{E}_t \mathbf{E}_t^\top)\|_2 \leq c p_2$，因此，

$$\begin{aligned}\|\mathbb{E}(\mathbf{E}_t^\top \mathbf{W}_1 \mathbf{W}_1^\top \mathbf{E}_t)\|_2 &\leq \mathrm{tr}(\mathbb{E}\mathbf{E}_t^\top \mathbf{W}_1 \mathbf{W}_1^\top \mathbf{E}_t) \leq \mathbb{E}(\mathrm{tr}(\mathbf{E}_t^\top \mathbf{W}_1 \mathbf{W}_1^\top \mathbf{E}_t)) \\ &= \mathrm{tr}(\mathbf{W}_1^\top \mathbb{E}(\mathbf{E}_t \mathbf{E}_t^\top)\mathbf{W}_1) \leq m_1 \|\mathbf{W}_1\|_2^2 \|\mathbb{E}(\mathbf{E}_t \mathbf{E}_t^\top)\|_2 \leq c m_1 p_1 p_2,\end{aligned}$$

其中倒数第二个不等式推导如下：设 $\boldsymbol{v}_i$ 是 $\mathsf{W}_1^\top \mathbb{E}(\mathsf{E}_t \mathsf{E}_t^\top)\mathsf{W}_1$, 的第 i 个特征向量，则

$$\mathrm{tr}(\mathbf{W}_1^\top \mathbb{E}(\mathbf{E}_t \mathbf{E}_t^\top)\mathbf{W}_1) = \sum_{i=1}^{m_1} \boldsymbol{v}_i^\top \mathbf{W}_1^\top \mathbb{E}(\mathbf{E}_t \mathbf{E}_t^\top)\mathbf{W}_1 \boldsymbol{v}_i \leq \|\mathbb{E}(\mathbf{E}_t \mathbf{E}_t^\top)\|_2 \sum_{i=1}^{m_1} \|\mathbf{W}_1 \boldsymbol{v}_i\|_2^2 \leq m_1 \|\mathbb{E}(\mathbf{E}_t \mathbf{E}_t^\top)\|_2 \|\mathbf{W}_1\|_2^2.$$

$$\begin{aligned}\|\mathbb{E}(\boldsymbol{\mathcal{E}}_t^\top \boldsymbol{\mathcal{E}}_t)\|_2 &= \frac{1}{p_1^2 p_2^2}\|\mathbb{E}(\mathbf{W}_2^\top \mathbf{E}_t^\top \mathbf{W}_1 \mathbf{W}_1^\top \mathbf{E}_t \mathbf{W}_2)\|_2 = \frac{1}{p_1^2 p_2^2}\|\mathbf{W}_2^\top \mathbb{E}(\mathbf{E}_t^\top \mathbf{W}_1 \mathbf{W}_1^\top \mathbf{E}_t)\mathbf{W}_2\|_2 \\ &\leq \frac{1}{p_1^2 p_2^2}\|\mathbf{W}_2\|_2^2 \|\mathbb{E}(\mathbf{E}_t^\top \mathbf{W}_1 \mathbf{W}_1^\top \mathbf{E}_t)\|_2,\end{aligned}$$

这意味着

$$\|\mathbb{E}(\boldsymbol{\mathcal{E}}_t^\top \boldsymbol{\mathcal{E}}_t)\|_2 \leq \frac{c}{p_1}.$$

因此，我们有 $\mathbb{E}\|\mathsf{E}_t\|_2^2 \leq \mathsf{tr}(\mathbb{E}\mathsf{E}_t^\top \mathsf{E}_t) \leq m_2 \|\mathbb{E}\mathsf{E}_t^\top \mathsf{E}_t\|_2 \leq \frac{c m_1 m_2}{p_1}$.

（2）成立：$\boldsymbol{\mathcal{E}} = \frac{1}{p_1 p_2}\mathbf{E}\mathbf{W}$，$\mathbf{W} = \mathbf{W}_2 \otimes \mathbf{W}_1$. 根据假设2.3,

$$\begin{aligned}\left\|\frac{1}{T}\mathbb{E}(\mathbf{E}^\top \mathbf{E})\right\|_2 &= \left\|\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\mathrm{Vec}(\mathbf{E}_t)\mathrm{Vec}(\mathbf{E}_t)^\top\right\|_2 \\ &= \|\mathbb{E}\mathrm{Vec}(\mathbf{E}_t)\mathrm{Vec}(\mathbf{E}_t)^\top\|_2 \\ &\leq \mathbb{E}\|\mathbb{E}\left(\mathrm{Vec}(\mathbf{E}_t)\mathrm{Vec}(\mathbf{E}_t)^\top|\mathbf{F}\right)\|_2 \leq c.\end{aligned}$$

Thus

$$\mathbb{E}\left\|\frac{1}{p_1 p_2}\mathbf{E}\mathbf{W}\right\|_2^2 \leq \frac{1}{p_1^2 p_2^2}\mathrm{tr}\mathbb{E}(\mathbf{W}^\top \mathbf{E}^\top \mathbf{E}\mathbf{W}) \leq \frac{1}{p_1^2 p_2^2}\|\mathbf{W}\|_F^2 \|\mathbb{E}\mathbf{E}^\top \mathbf{E}\|_2 \leq \frac{cT}{p_1 p_2}.$$

（3）根据假设2.3，可得

$$\frac{1}{T}\sum_{s=1}^{T}\sum_{T=1}^{T}\mathbb{E}\|\mathbf{F}_t\|_F\|\mathbf{F}_s\|_F\|\mathbb{E}(\mathrm{Vec}(\mathbf{E}_s)\mathrm{Vec}(\mathbf{E}_t)^\top|\mathbf{F})\|_2 < c,$$

那么我们可以得到

$$\begin{aligned}
&\mathbb{E}\|\frac{1}{T}\mathbf{F}^\top\boldsymbol{\mathcal{E}}\|_2^2 = \frac{1}{T^2p_1^2p_2^2}\mathbb{E}\|\mathbf{F}^\top\mathbf{E}\mathbf{W}\|_2^2 = \frac{1}{T^2p_1^2p_2^2}\mathbb{E}\|\sum_{t=1}^{T}\mathbf{W}^\top\mathrm{Vec}(\mathbf{E}_t)\mathrm{Vec}(\mathbf{F}_t)^\top\|_2^2\\
&= \frac{1}{T^2p_1^2p_2^2}\mathbb{E}\left[\mathbb{E}\left(\|\sum_{t=1}^{T}\mathbf{W}^\top\mathrm{Vec}(\mathbf{E}_t)\mathrm{Vec}(\mathbf{F}_t)^\top\|_2^2\right)\middle|\mathbf{F}\right]\\
&\le \frac{1}{T^2p_1^2p_2^2}\mathbb{E}\left[\mathbb{E}\left(\mathrm{tr}(\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{Vec}(\mathbf{F}_t)\mathrm{Vec}(\mathbf{E}_t)^\top\mathbf{W}\mathbf{W}^\top\mathrm{Vec}(\mathbf{E}_s)\mathrm{Vec}(\mathbf{F}_s)^\top)\right)\middle|\mathbf{F}\right]\\
&= \frac{1}{T^2p_1^2p_2^2}\mathbb{E}\left\{\mathrm{tr}\left[\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{Vec}(\mathbf{F}_t)\mathbb{E}\left(\mathrm{Vec}(\mathbf{E}_t)^\top\mathbf{W}\mathbf{W}^\top\mathrm{Vec}(\mathbf{E}_s)\middle|\mathbf{F}\right)\mathrm{Vec}(\mathbf{F}_s)^\top\right]\right\}\\
&= \frac{1}{T^2p_1^2p_2^2}\mathbb{E}\left\{\mathrm{tr}\left[\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{Vec}(\mathbf{F}_t)\mathrm{Vec}(\mathbf{F}_s)^\top\mathbb{E}\left(\mathrm{Vec}(\mathbf{E}_t)^\top\mathbf{W}\mathbf{W}^\top\mathrm{Vec}(\mathbf{E}_s)\middle|\mathbf{F}\right)\right]\right\}\\
&\le \frac{1}{T^2p_1^2p_2^2}\sum_{i=1}^{k_1}\sum_{j=1}^{k_2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}\left[f_{t,ij}f_{s,ij}\mathbb{E}\left(\mathrm{Vec}(\mathbf{E}_t)^\top\mathbf{W}\mathbf{W}^\top\mathrm{Vec}(\mathbf{E}_s)\middle|\mathbf{F}\right)\right]\\
&= \frac{1}{T^2p_1^2p_2^2}\sum_{i=1}^{k_1}\sum_{j=1}^{k_2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}\left\{f_{t,ij}f_{s,ij}\mathrm{tr}\left[\mathbf{W}^\top\mathbb{E}\left(\mathrm{Vec}(\mathbf{E}_s)\mathrm{Vec}(\mathbf{E}_t)^\top\middle|\mathbf{F}\right)\mathbf{W}\right]\right\}\\
&\le \frac{1}{T^2p_1^2p_2^2}\sum_{i=1}^{k_1}\sum_{j=1}^{k_2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}\left[|f_{t,ij}f_{s,ij}|\|\mathbf{W}\|_F^2\left\|\mathbb{E}\left(\mathrm{Vec}(\mathbf{E}_s)\mathrm{Vec}(\mathbf{E}_t)^\top\middle|\mathbf{F}\right)\right\|_2\right]\\
&\le \frac{c}{T^2p_1p_2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}\|\mathbf{F}_t\|_F\|\mathbf{F}_s\|_F\left\|\mathbb{E}\left(\mathrm{Vec}(\mathbf{E}_s)\mathrm{Vec}(\mathbf{E}_t)^\top\middle|\mathbf{F}\right)\right\|_2 \le \frac{c}{Tp_1p_2}.
\end{aligned}$$

因此， $\mathbb{E}\|\frac{1}{T}\mathbf{F}^\top\mathbf{E}$
$\|_2^2 \le O\left(\frac{1}{Tp_1p_2}\right)$.

（4）根据假设2.3 （4），我们有

$$
\begin{aligned}
&\mathbb{E}\|\frac{1}{T}(\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}}-\mathbb{E}\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}})\|_2^2 \le \mathbb{E}\|\frac{1}{T}(\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}}-\mathbb{E}\boldsymbol{\mathcal{E}}^\top\boldsymbol{\mathcal{E}})\|_F^2 \\
&\le \sum_{k,q=1}^{m_1m_2}\mathbb{E}\left(\frac{1}{Tp_1^2p_2^2}\sum_{t=1}^{T}\sum_{i,j=1}^{p_1p_2}w_{ik}w_{jq}(e_{ti}e_{tj}-\mathbb{E}e_{ti}e_{tj})\right)^2 \\
&= \sum_{k,q=1}^{m_1m_2}\mathbb{E}\left(\frac{1}{Tp_1^2p_2^2}\sum_{t=1}^{T}\sum_{i,j=1}^{p_1p_2}w_{ik}w_{jq}\left(\mathrm{Vec}(\mathbf{E}_t)_i\mathrm{Vec}(\mathbf{E}_t)_j-\mathbb{E}\mathrm{Vec}(\mathbf{E}_t)_i\mathrm{Vec}(\mathbf{E}_t)_j\right)\right)^2 \\
&\le \frac{1}{Tp_1^2p_2^2}\sum_{k,q=1}^{m_1m_2}\frac{1}{Tp_1^2p_2^2}\sum_{t,s=1}^{T}\sum_{i,j,u,v=1}^{p_1p_2}\left|w_{ik}w_{jq}w_{uk}w_{vq}\right|\left|\mathrm{Cov}\left(\mathrm{Vec}(\mathbf{E}_t)_i\mathrm{Vec}(\mathbf{E}_t)_j,\mathrm{Vec}(\mathbf{E}_s)_u\mathrm{Vec}(\mathbf{E}_s)_v\right)\right| \\
&\le \frac{c}{Tp_1^2p_2^2}\frac{1}{Tp_1^2p_2^2}\sum_{t,s=1}^{T}\sum_{i_1,j_1,u_1,v_1=1}^{p_1}\sum_{i_2,j_2,u_2,v_2=1}^{p_2}|\mathrm{Cov}(e_{t,i_1i_2}e_{t,j_1j_2},e_{s,u_1u_2}e_{s,v_1v_2})| \le \frac{c}{Tp_1^2p_2^2}.
\end{aligned}
$$

□

引理 S2.3. 在假设 2.1 (1) 和假设 3.2-3.4成立时，当 $\min\{T,p_1,p_2\}\to\infty$时，我们有 (1) $\sum_{t=1}^{T}\mathbb{E}\|\mathbf{E}_t^\top\mathbf{R}\|_F^2=O\left(Tp_1p_2\right)$, $\sum_{t=1}^{T}\mathbb{E}\|\mathbf{E}_t\mathbf{C}\|_F^2=O\left(Tp_1p_2\right)$, $\mathbb{E}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{F}_t\|_F^2=O\left(Tp_1p_2\right)$,

(2) $\mathbb{E}\|\sum_{t=1}^{T}\mathrm{F}_t\mathrm{C}^\top\mathrm{E}_t^\top\|_F^2=O\left(Tp_1p_2\right)$ $\mathbb{E}\|\sum_{t=1}^{T}\mathrm{F}_t^\top\mathrm{R}^\top\mathrm{E}_t\|_F^2=O\left(Tp_1p_2\right)$,

$\mathbb{E}\|\sum_{t=1}^{T}\mathrm{F}_t\mathrm{C}^\top\mathrm{E}_t^\top\mathrm{R}\|_F^2=O\left(Tp_1p_2\right)$ $\mathbb{E}\|\sum_{t=1}^{T}\mathrm{F}_t^\top\mathrm{R}^\top\mathrm{E}_t\mathrm{C}\|_F^2=O\left(Tp_1p_2\right)$,

(3) $\mathbb{E}\|\sum_{t=1}^{T}\mathrm{F}_t\mathrm{W}_2^\top\mathrm{E}_t^\top\mathrm{W}_1\|_F^2=O\left(Tp_1p_2\right)$ $\mathbb{E}\|\sum_{t=1}^{T}\mathrm{E}_t^\top\mathrm{W}_1\mathrm{F}_t\|_F^2=O\left(Tp_1p_2\right)$,

$\mathbb{E}\|\sum_{t=1}^{T}\mathrm{F}_t\mathrm{W}_2^\top\mathrm{E}_t^\top\|_F^2=O\left(Tp_1p_2\right)$ $\mathbb{E}\|\sum_{t=1}^{T}\mathrm{W}_1^\top\mathrm{E}_t\mathrm{W}_2\|_F^2=O\left(Tp_1p_2\right)$,

(4) $\mathbb{E}\|\sum_{t=1}^{T}\mathrm{R}^\top\mathrm{E}_t\mathrm{W}_2\mathrm{F}_t^\top\|_F^2=O\left(Tp_1p_2\right)$ $\mathbb{E}\|\sum_{t=1}^{T}\mathrm{R}^\top\mathrm{E}_t\mathrm{W}_2\|_F^2=O\left(Tp_1p_2\right)$

(5) $\mathbb{E}\|\sum_{t=1}^{T}\text{东}_t\text{西}_2\text{西}_2^\top\text{东}_t^\top\text{西}_1\|_F^2=O\left(Tp_1^2p_2^3+T^2p_1p_2^2\right)$。

证明。假设 $m_1=k_1=1,m_2=k_2=1$。

(1) 根据假设 3.2，我们可以得到

$$
\mathbb{E}\|\mathbf{E}_t^\top\mathbf{R}\|_F^2=\sum_{j=1}^{p_2}\mathbb{E}(\sum_{i=1}^{p_1}e_{t,ij}r_i)^2=\sum_{j=1}^{p_2}\sum_{i_1,i_2=1}^{p_1}\mathbb{E}(r_{i_1}r_{i_2}e_{t,i_1j}e_{t,i_2j})=O\left(p_1p_2\right).
$$

类似地，$\sum_{t=1}^{T}\mathbb{E}\|\mathbf{E}_t\mathbf{C}\|_F^2=O\left(Tp_1p_2\right)$。根据假设 3.4(2),

$$
\mathbb{E}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{F}_t\|_F^2=\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\mathbb{E}(\sum_{t=1}^{T}\mathbf{F}_te_{t,ij})^2=T\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\mathbb{E}(\boldsymbol{\xi}_{i,j})^2=O\left(Tp_1p_2\right).
$$

(2) 由假设 3.4 (1)，结论直接成立。

（3）根据假设2.1 （1）和假设3.4 （2），可得

$$\mathbb{E}\|\sum_{t=1}^{T}\mathbf{F}_t\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\|_F^2=\mathbb{E}(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\mathbf{F}_t w_{1,i}e_{t,ij}w_{2,j})^2=T\mathbb{E}(\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\boldsymbol{\xi}_{i,j}w_{1,i}w_{2,j})^2$$
$$=T\sum_{i_1,i_2=1}^{p_1}\sum_{j_1,j_2=1}^{p_2}\mathbb{E}(\boldsymbol{\xi}_{i_1,j_1}\boldsymbol{\xi}_{i_2,j_2})w_{1,i_1}w_{1,i_2}w_{2,j_1}w_{2,j_2}=O\left(Tp_1p_2\right).$$

$$\mathbb{E}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{W}_1\mathbf{F}_t\|_F^2=\sum_{j=1}^{p_2}\mathbb{E}(\sum_{t=1}^{T}\sum_{i=1}^{p_1}e_{t,ij}w_{1,j}\mathbf{F}_t)^2=T\sum_{j=1}^{p_2}\mathbb{E}(\sum_{i=1}^{p_1}\boldsymbol{\xi}_{i,j}w_{1,j})^2$$
$$=T\sum_{j=1}^{p_2}\sum_{i,i'=1}^{p_1}\mathbb{E}(\boldsymbol{\xi}_{i,j}\boldsymbol{\xi}_{i',j})w_{1,i}w_{1,i'}=O\left(Tp_1p_2\right).$$

类似地， $\mathbb{E}\|\sum_{t=1}^{T}\mathbf{F}_t\mathbf{W}_2^\top\mathbf{E}_t^\top\|_F^2=O(Tp_1p_2)$ 。根据假设3.2 （2），我们也可以得到

$$\mathbb{E}\|\sum_{t=1}^{T}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\|_F^2=\mathbb{E}(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}w_{1,i}w_{2,j}e_{t,ij})^2=\sum_{s,t=1}^{T}\sum_{i_1,i_2=1}^{p_1}\sum_{j_1,j_2=1}^{p_2}w_{1,i_1}w_{1,i_2}w_{2,j_1}w_{2,j_2}\mathbb{E}(e_{t,i_1j_1}e_{s,i_2j_2})^2$$
$$=O\left(Tp_1p_2\right).$$

(4)

$$\mathbb{E}\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\mathbf{W}_2\mathbf{F}_t\|_F^2=\mathbb{E}(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}r_iw_{2,j}e_{t,ij}\mathbf{F}_t)^2=T\mathbb{E}(\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}r_iw_{2,j}\boldsymbol{\xi}_{i,j})^2$$
$$=T\sum_{i_1,i_2=1}^{p_1}\sum_{j_1,j_2=1}^{p_2}r_{i_1}r_{i_2}w_{2,j_1}w_{2,j_2}\mathbb{E}(\boldsymbol{\xi}_{i_1,j_1}\boldsymbol{\xi}_{i_2,j_2})=O\left(Tp_1p_2\right).$$

$$\mathbb{E}\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\mathbf{W}_2\|_F^2=\mathbb{E}(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}r_iw_{2,j}e_{t,ij})^2=\sum_{s,t=1}^{T}\sum_{i_1,i_2=1}^{p_1}\sum_{j_1,j_2=1}^{p_2}r_{i_1}r_{i_2}w_{2,j_1}w_{2,j_2}\mathbb{E}(e_{t,i_1j_1}e_{s,i_2j_2})^2$$
$$=O\left(Tp_1p_2\right).$$

（5）注意,

$$\mathbb{E}\|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\|_F^2\le\|\mathbf{W}_2\|_F^2\|\sum_{t=1}^{T}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\mathbf{E}_t\|_F^2\le\|\mathbf{W}_2\|_F^2\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\mathbb{E}\|\sum_{t=1}^{T}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2e_{t,ij}\|^2,$$

对于任意 i， j根据假设 3.2,

$$\mathbb{E}\|\sum_{t=1}^{T}\boldsymbol{W}_1^\top\mathbf{E}_t\mathbf{W}_2e_{t,ij}\|^2=\mathbb{E}(\sum_{t=1}^{T}\sum_{i'=1}^{p_1}\sum_{j'=1}^{p_2}w_{1,i'}e_{t,i'j'}w_{2,j'}e_{t,ij})^2$$
$$\le c\sum_{t,s=1}^{T}\sum_{i_1,i_2=1}^{p_1}\sum_{j_1,j_2=1}^{p_2}|\mathrm{Cov}(e_{t,ij}e_{t,i_1j_1},e_{s,ij}e_{s,i_2j_2})|+c(\sum_{t=1}^{T}\sum_{i'=1}^{p_1}\sum_{j'=1}^{p_2}|\mathbb{E}e_{t,i'j'}e_{t,ij}|)^2$$
$$=O\left(Tp_1p_2+T^2\right).$$

□

引理 S2.4。 在假设 2.1 (1)、假设 2.2 以及假设 3.1-3.4，当 $\min\{T, p_1, p_2\} \to \infty$时，我们有

$$\|\sum_{s=1}^{T} \mathbf{E}_s^\top (\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{F}_s\|_F^2 = O_p\left(\frac{p_1^2}{p_2\nu_{\min}^2(\mathbf{H}_2)} + \frac{p_1}{\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)} + \frac{T}{p_2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}\right).$$

证明。 根据定理 3.1， $\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)} = \mathcal{I}\mathcal{I} + \mathcal{I}\mathcal{I}\mathcal{I} + \mathcal{I}\mathcal{V}$，则

$$\sum_{s=1}^{T} \mathbf{E}_s^\top (\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{F}_s = \sum_{s=1}^{T} \mathbf{E}_s^\top (\mathcal{I}\mathcal{I} + \mathcal{I}\mathcal{I}\mathcal{I} + \mathcal{I}\mathcal{V})\mathbf{F}_s.$$

对于第一项，根据引理S2.3 （2）和（3），可得

$$\begin{aligned}
\|\sum_{t=1}^{T} \mathbf{E}_s^\top \mathcal{I}\mathcal{I}\mathbf{F}_s\|_F^2 &= \frac{1}{T^2p_1p_2^2}\|\sum_{s=1}^{T} \mathbf{E}_s^\top \mathbf{R}[(\sum_{t=1}^{T} \mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}]\mathbf{F}_s\|_F^2 \\
&\le \frac{1}{T^2p_1p_2^2}\|\sum_{s=1}^{T} \mathbf{E}_s^\top \mathbf{R}\mathbf{F}_s\|_F^2\|(\sum_{t=1}^{T} \mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}\|_F^2 \\
&\le \frac{1}{T^2p_1p_2^2}\|\sum_{s=1}^{T} \mathbf{E}_s^\top \mathbf{R}\mathbf{F}_s\|_F^2\|\sum_{t=1}^{T} \mathbf{F}_t\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\|_F^2\|\mathbf{Z}^{-1/2}\|_F^2\|\mathbf{C}^\top\mathbf{W}_2\|_2^2 \\
&= O_p\left(\frac{1}{\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}\right).
\end{aligned}$$

对于第二项，

$$\begin{aligned}
\|\sum_{s=1}^{T} \mathbf{E}_s^\top \mathcal{I}\mathcal{I}\mathcal{I}\mathbf{F}_s\|_F^2 &= \frac{1}{T^2p_1p_2^2}\|\sum_{s=1}^{T} \mathbf{E}_s^\top [(\sum_{t=1}^{T} \mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}]\mathbf{F}_s\|_F^2 \\
&\le \frac{1}{T^2p_1p_2^2}\|\sum_{s=1}^{T} \mathbf{E}_s^\top (\sum_{t=1}^{T} \mathbf{E}_t\mathbf{W}_2\mathbf{F}_t^\top)\mathbf{F}_s\|_F^2\|\mathbf{W}_2^\top\mathbf{C}\|_F^2\|\mathbf{R}^\top\mathbf{W}_1\|_F^2\|\mathbf{Z}^{-1/2}\|_F^2 \\
&= O_p\left(\frac{p_1^2}{p_2\nu_{\min}^2(\mathbf{H}_2)} + \frac{p_1}{\nu_{\min}^2(\mathbf{H}_2)}\right),
\end{aligned}$$

其中

$$\begin{aligned}
\mathbb{E}\|\sum_{s=1}^{T} \mathbf{E}_s^\top (\sum_{t=1}^{T} \mathbf{E}_t\mathbf{W}_2\mathbf{F}_t^\top)\mathbf{F}_s\|_F^2 &= \sum_{j=1}^{p_2}\mathbb{E}(\sum_{s,t=1}^{T}\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2} w_{2,j_1}\mathbf{F}_s\mathbf{F}_t e_{s,ij}e_{t,ij_1})^2 = T^2\sum_{j=1}^{p_2}\mathbb{E}(\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2} w_{2,j_1}\boldsymbol{\xi}_{i,j}\boldsymbol{\xi}_{i,j_1})^2 \\
&= T^2\sum_{j=1}^{p_2}\left[(\sum_{i=1}^{p_1}\sum_{j_1=1}^{p_2}\mathbb{E}\boldsymbol{\xi}_{i,j}\boldsymbol{\xi}_{i,j_1})^2 + \sum_{i_1,i_2=1}^{p_1}\sum_{j_1j_2=1}^{p_2}\mathrm{Cov}(\boldsymbol{\xi}_{i_1,j}\boldsymbol{\xi}_{i_1,j_1}, \boldsymbol{\xi}_{i_2,j}\boldsymbol{\xi}_{i_2,j_2})\right] \\
&= O\left(T^2p_1^2p_2 + T^2p_1p_2^2\right).
\end{aligned}$$

对于第三项，由引理 S2.3 （1）和（5），可得

$$\begin{aligned}\|\sum_{s=1}^{T}\mathbf{E}_s^\top\mathcal{IV}\mathbf{F}_s\|_F^2 &= \frac{1}{T^2p_1p_2^2}\|\sum_{s=1}^{T}\mathbf{E}_s[(\sum_{t=1}^{T}\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}]\mathbf{F}_s\|_F^2\\ &\le \frac{1}{T^2p_1p_2^2}\|\sum_{s=1}^{T}\mathbf{E}_s^\top\mathbf{F}_s\|_F^2\|(\sum_{t=1}^{T}\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}\|_F^2\\ &= O_p\left(\frac{p_1}{\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}+\frac{T}{p_2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}\right).\end{aligned}$$

因此，

$$\|\sum_{s=1}^{T}\mathbf{E}_s^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{F}_s\|_F^2 = O_p\left(\frac{p_1^2}{p_2\nu_{\min}^2(\mathbf{H}_2)}+\frac{p_1}{\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}+\frac{T}{p_2\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}\right).$$

□

引理 S2.5。 在假设 2.1 (1)、假设 2.2 和假设 3.1-3.4成立的条件下，取 $\widehat{\ }\mathbf{R}^{(1)}$, $\widehat{\ }\mathbf{C}^{(1)}$ 为单步迭代所得的合成估计量，当 $\min\{T,p_1,p_2\}\to\infty$时，我们有

$$\left\|\mathbf{R}^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\right\|_F = O_p\left(\frac{\sqrt{p_1}}{\sqrt{Tp_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}+\frac{1}{p_2\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}\right),$$

$$\left\|\mathbf{C}^\top(\widehat{\mathbf{C}}^{(1)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(1)})\right\|_F = O_p\left(\frac{1}{T\nu_{\min}(\mathbf{H}_2)}+\frac{\sqrt{p_2}}{\sqrt{Tp_1}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}+\frac{1}{p_1\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}\right).$$

证明。

$$\begin{aligned}\|\mathbf{R}^\top\mathcal{II}\|_F &= \frac{1}{T\sqrt{p_1}p_2}\left\|\mathbf{R}^\top\mathbf{R}(\sum_{t=1}^{T}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}\right\|_F = \frac{\sqrt{p_1}}{Tp_2}\left\|(\sum_{t=1}^{T}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}\right\|_F\\ &\le \frac{\sqrt{p_1}}{Tp_2}\left\|\sum_{t=1}^{T}\mathbf{F}_t\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\right\|_F\left\|\mathbf{Z}^{-1/2}\right\|_F\|\mathbf{C}^\top\mathbf{W}_2\|_F = O_p\left(\frac{\sqrt{p_1}}{\sqrt{Tp_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}\right),\end{aligned}$$

$$\begin{aligned}\|\mathbf{R}^\top\mathcal{III}\|_F &= \frac{1}{T\sqrt{p_1}p_2}\left\|(\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{C}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}\right\|_F\\ &\le \frac{1}{T\sqrt{p_1}p_2}\left\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\mathbf{W}_2\mathbf{F}_t^\top\right\|_F\|\mathbf{W}_2^\top\mathbf{C}\|_F\|\mathbf{R}^\top\mathbf{W}_1\|_F\left\|\mathbf{Z}^{-1/2}\right\|_F\\ &= O_p\left(\frac{\sqrt{p_1}}{\sqrt{Tp_2}\nu_{\min}(\mathbf{H}_2)}\right),\end{aligned}$$

$$\begin{aligned}\|\mathbf{R}^\top\mathcal{IV}\|_F &= \frac{1}{T\sqrt{p_1}p_2}\left\|(\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\mathbf{W}_2\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1)\mathbf{Z}^{-1/2}\right\|_F\\ &\le \frac{1}{T\sqrt{p_1}p_2}\sqrt{\left\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\mathbf{W}_2\right\|_F^2\times\left\|\sum_{t=1}^{T}\mathbf{W}_2^\top\mathbf{E}_t^\top\mathbf{W}_1\right\|_F^2}\left\|\mathbf{Z}^{-1/2}\right\|_F\\ &= O_p\left(\frac{1}{p_2\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}\right).\end{aligned}$$

由于 $\mathrm{R}^\top(\widehat{\mathrm{R}}^{(1)} - \mathrm{R}\widehat{\mathrm{H}}_r^{(1)}) = \mathrm{R}^\top(\mathcal{II} + \mathcal{III} + \mathcal{IV})$，综合以上三项，我们得到

$$\left\|\mathbf{R}^\top(\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\right\|_F = O_p\left(\frac{\sqrt{p_1}}{\sqrt{Tp_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)} + \frac{1}{p_2\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}\right).$$

对于第二个公式，由定理 3.2 的证明可得

$$\widehat{\mathbf{C}}^{(1)} - \mathbf{C}\widehat{\mathbf{H}}_c^{(1)} = \sqrt{p_2}\left(\boldsymbol{\delta}_2^{(1)}(\mathbf{Z}^{(1)})^{-1/2} + \boldsymbol{\delta}_3^{(1)}(\mathbf{Z}^{(1)})^{-1/2} + \boldsymbol{\delta}_4^{(1)}(\mathbf{Z}^{(1)})^{-1/2}\right).$$

对于第一项，

$$\begin{aligned}
\sqrt{p_2}\left\|\mathbf{C}^\top\boldsymbol{\delta}_2^{(1)}(\mathbf{Z}^{(1)})^{-1/2}\right\|_F &= \frac{1}{p_1\sqrt{p_2}}\left\|(\sum_{t=1}^T \mathbf{C}^\top\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{R}\mathbf{F}_t\mathbf{C}^\top\mathbf{W}_2)(\mathbf{Z}^{(1)})^{-1/2}\right\|_F \\
&\le \frac{1}{p_1\sqrt{p_2}}\left\|\sum_{t=1}^T \mathbf{C}^\top\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{F}_t\right\|_F \left\|\mathbf{W}_1^\top\mathbf{R}\right\|_2 \left\|\mathbf{C}^\top\mathbf{W}_2\right\|_2 \left\|(\mathbf{Z}_t^{(1)})^{-1/2}\right\|_F \\
&= O_p\left(\frac{1}{T\nu_{\min}(\mathbf{H}_2)} + \frac{1}{T\sqrt{p_1}\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)} + \frac{\sqrt{p_2}}{\sqrt{Tp_1}}\right),
\end{aligned}$$

其中

$$\begin{aligned}
\left\|\sum_{t=1}^T \mathbf{C}^\top\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{F}_t\right\|_F &\le \left\|\sum_{t=1}^T \mathbf{C}^\top\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\mathbf{F}_t\right\|_F + \left\|\sum_{t=1}^T \mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\mathbf{F}_t\right\|_F \\
&\le \left\|\sum_{t=1}^T \mathbf{C}^\top\mathbf{E}_t^\top\mathbf{F}_t\right\|_F \left\|\widehat{\mathbf{R}}^{(1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\right\|_F + \left\|\sum_{t=1}^T \mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\mathbf{F}_t\right\|_F \left\|\widehat{\mathbf{H}}_r^{(1)}\right\|_F \\
&= O_p\left(\frac{p_1}{\nu_{\min}(\mathbf{H}_2)} + \frac{\sqrt{p_1}}{\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)} + \sqrt{Tp_1p_2}\right).
\end{aligned}$$

对于第二项，

$$\begin{aligned}
\sqrt{p_2}\left\|\mathbf{C}^\top\boldsymbol{\delta}_3^{(1)}(\mathbf{Z}^{(1)})^{-1/2}\right\|_F &= \frac{\sqrt{p_2}}{p_1}\left\|(\sum_{t=1}^T \mathbf{F}_t^\top\mathbf{R}^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2)(\mathbf{Z}^{(1)})^{-1/2}\right\|_F \\
&\asymp \sqrt{p_2}\left\|(\sum_{t=1}^T \mathbf{F}_t^\top\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2)(\mathbf{Z}^{(1)})^{-1/2}\right\|_F \\
&\le \sqrt{p_2}\left\|\sum_{t=1}^T \mathbf{F}_t^\top\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F \left\|(\mathbf{Z}^{(1)})^{-1/2}\right\|_F \\
&= O_p\left(\frac{\sqrt{p_2}}{\sqrt{Tp_1}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}\right).
\end{aligned}$$

对于第三项,

$$\begin{aligned}
&\sqrt{p_2}\left\|\mathbf{C}^\top\boldsymbol{\delta}_4^{(1)}(\mathbf{Z}^{(1)})^{-1/2}\right\|_F = \frac{1}{p_1\sqrt{p_2}}\left\|(\sum_{t=1}^T\mathbf{C}^\top\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2)(\mathbf{Z}^{(1)})^{-1/2}\right\|_F \\
&\le \frac{1}{p_1\sqrt{p_2}}\sqrt{\left\|\sum_{t=1}^T\mathbf{C}^\top\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\right\|_F^2\times\left\|\sum_{t=1}^T\mathbf{W}_1^\top\mathbf{E}_t\mathbf{W}_2\right\|_F^2}\left\|(\mathbf{Z}^{(1)})^{-1/2}\right\|_F \\
&= O_p\left(\frac{1}{\sqrt{Tp_1p_2}\nu_{\min}(\mathbf{H}_1)\nu_{\min}^2(\mathbf{H}_2)}+\frac{1}{\sqrt{Tp_2}p_1\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^3(\mathbf{H}_2)}+\frac{1}{p_1\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}\right),
\end{aligned}$$

其中，最后一个等式成立是因为

$$\begin{aligned}
&\left\|\sum_{t=1}^T\mathbf{C}^\top\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(1)}\right\|_F^2 = \left\|\sum_{t=1}^T\mathbf{C}^\top\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}+\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\right\|_F^2 \\
&\le \left\|\sum_{t=1}^T\mathbf{C}^\top\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(1)})\right\|_F^2+\left\|\sum_{t=1}^T\mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(1)}\right\|_F^2 \\
&= O_p\left(\frac{p_1^2}{\nu_{\min}^2(\mathbf{H}_2)}+\frac{p_1}{\nu_{\min}^2(\mathbf{H}_1)\nu_{\min}^4(\mathbf{H}_2)}+Tp_1p_2\right).
\end{aligned}$$

因此,

$$\left\|\mathbf{C}^\top(\widehat{\mathbf{C}}^{(1)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(1)})\right\|_F = O_p\left(\frac{1}{T\nu_{\min}(\mathbf{H}_2)}+\frac{\sqrt{p_2}}{\sqrt{Tp_1}\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}+\frac{1}{p_1\nu_{\min}(\mathbf{H}_1)\nu_{\min}(\mathbf{H}_2)}\right).$$

□

引理 S2.6。 在假设 2.1 (1)、假设 2.2 和假设 3.1-3.4下，取 $\widehat{\ }\mathbf{R}^{(s+1)}$, $\widehat{\ }\mathbf{C}^{(s+1)}$ 作为第（$s+1$）步迭代的结果，其中 $\min\{T,p_1,p_2\}\to\infty$，则有

$$\|\mathbf{R}^\top(\widehat{\mathbf{R}}^{(s+1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)})\|_F = O_p\left(\frac{\sqrt{p_1}}{\sqrt{Tp_2}}+\frac{p_1\sqrt{w_r^{(s)}}}{\sqrt{Tp_2}}+\frac{\sqrt{w_r^{(s)}}}{p_2}+\frac{\sqrt{p_1w_c^{(s)}}}{\sqrt{T}}+\frac{p_1\sqrt{w_r^{(s)}w_c^{(s)}}}{\sqrt{T}}+\sqrt{w_r^{(s)}}w_c^{(s)}\right),$$

$$\|\mathbf{C}^\top(\widehat{\mathbf{C}}^{(s+1)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s+1)})\|_F = O_p\left(\frac{\sqrt{p_2}}{\sqrt{Tp_1}}+\frac{1}{p_1}+\frac{\sqrt{w_r^{(s)}}}{\sqrt{p_1}}+\frac{\sqrt{p_2w_r^{(s)}}}{\sqrt{T}}+\frac{p_2\sqrt{w_c^{(s)}}}{\sqrt{Tp_1}}+\frac{p_2\sqrt{w_r^{(s)}w_c^{(s)}}}{\sqrt{T}}+\sqrt{w_r^{(s)}w_r^{(s+1)}}\right).$$

证明。 根据定理 3.4,

$$\widehat{\mathbf{R}}^{(s+1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)} = \sqrt{p_1}(\boldsymbol{\delta}_2^{(s+1)}+\boldsymbol{\delta}_3^{(s+1)}+\boldsymbol{\delta}_4^{(s+1)})(\mathbf{Z}^{(s+1)})^{-1/2}.$$

根据引理 S2.9 (1) 和 (2)，我们有

$$\begin{aligned}\left\|\sqrt{p_1}\mathbf{R}^\top\boldsymbol{\delta}_2^{(s+1)}(\mathbf{Z}^{(s+1)})^{-1/2}\right\|_F &\asymp \left\|\sqrt{p_1}\left(\sum_{t=1}^T\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\right\|_F\\ &\le \sqrt{p_1}\left\|\sum_{t=1}^T\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\right\|_F\left\|\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\right\|_F\\ &= O_p\left(\frac{p_1\sqrt{w_r^{(s)}w_c^{(s)}}}{\sqrt{T}}+\frac{p_1\sqrt{w_r^{(s)}}}{\sqrt{Tp_2}}+\frac{\sqrt{p_1w_c^{(s)}}}{\sqrt{T}}+\frac{\sqrt{p_1}}{\sqrt{Tp_2}}\right),\end{aligned}$$

其中，最后一个等式成立是因为

$$\left\|\sum_{t=1}^T\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\right\|_F \le \left\|\sum_{t=1}^T\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\right\|_F\left\|\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\right\|_F+\left\|\sum_{t=1}^T\mathbf{F}_t\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\mathbf{R}\right\|_F\left\|\widehat{\mathbf{H}}_r^{(s)}\right\|_F.$$

$$\begin{aligned}\left\|\sqrt{p_1}\mathbf{R}^\top\boldsymbol{\delta}_3^{(s+1)}(\mathbf{Z}^{(s+1)})^{-1/2}\right\|_F &\asymp \sqrt{p_1}\left\|\left(\sum_{t=1}^T\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\right\|_F\\ &\le \sqrt{p_1}\left\|\sum_{t=1}^T\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\right\|_F\left\|\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\right\|_F\\ &= O_p\left(\frac{\sqrt{p_1w_c^{(s)}}}{\sqrt{T}}+\frac{\sqrt{p_1}}{\sqrt{Tp_2}}\right).\end{aligned}$$

由.....这一事实可知

$$\begin{aligned}\left\|\sum_{t=1}^T\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\right\|_F &\le \left\|\sum_{t=1}^T\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)})\right\|_F+\left\|\sum_{t=1}^T\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\right\|_F\\ &\le \left\|\sum_{t=1}^T\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\right\|_F\left\|\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\right\|_F+\left\|\sum_{t=1}^T\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\right\|_F^2\left\|\widehat{\mathbf{H}}_r^{(s)}\right\|_F,\end{aligned}$$

以及引理 S2.8 (1)，可得

$$\begin{aligned}\left\|\sqrt{p_1}\mathbf{R}^\top\boldsymbol{\delta}_4^{(s+1)}(\mathbf{Z}^{(s+1)})^{-1/2}\right\|_F &= \left\|\frac{1}{\sqrt{p_1}p_2}\left(\sum_{t=1}^T\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\right)\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\right\|_F\\ &\le \frac{1}{\sqrt{p_1}p_2}\left\|\sum_{t=1}^T\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\right\|_F\left\|\left(\mathbf{Z}^{(s+1)}\right)^{-1/2}\right\|_F\\ &= O_p\left(\frac{\sqrt{p_1w_r^{(s)}}}{\sqrt{Tp_2}}+\frac{\sqrt{w_r^{(s)}}}{p_2}+\sqrt{w_r^{(s)}}w_c^{(s)}+\frac{\sqrt{p_1w_r^{(s)}}w_c^{(s)}}{\sqrt{T}}+\frac{\sqrt{w_c^{(s)}}}{\sqrt{Tp_1}p_2}+\frac{1}{\sqrt{Tp_1p_2}p_2}\right).\end{aligned}$$

因此，

$$\|\mathbf{R}^\top(\widehat{\mathbf{R}}^{(s+1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)})\|_F = O_p\left(\frac{\sqrt{p_1}}{\sqrt{Tp_2}}+\frac{p_1\sqrt{w_r^{(s)}}}{\sqrt{Tp_2}}+\frac{\sqrt{w_r^{(s)}}}{p_2}+\frac{\sqrt{p_1w_c^{(s)}}}{\sqrt{T}}+\frac{p_1\sqrt{w_r^{(s)}w_c^{(s)}}}{\sqrt{T}}+\sqrt{w_r^{(s)}}w_c^{(s)}\right).$$

通过类似的证明，我们可以得到

$$\|\mathbf{C}^\top(\widehat{\mathbf{C}}^{(s+1)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s+1)})\|_F = O_p\left(\frac{\sqrt{p_2}}{\sqrt{Tp_1}}+\frac{1}{p_1}+\frac{\sqrt{w_r^{(s)}}}{\sqrt{p_1}}+\frac{\sqrt{p_2 w_r^{(s)}}}{\sqrt{T}}+\frac{p_2\sqrt{w_c^{(s)}}}{\sqrt{Tp_1}}+\frac{p_2\sqrt{w_r^{(s)}w_c^{(s)}}}{\sqrt{T}}+\sqrt{w_r^{(s)}w_r^{(s+1)}}\right).$$

□

引理 S2.7. 假设 T、p_1、p_2趋于无穷大，且 $m_1=k_1$、$m_2=k_2$为固定值。若假设 2.1(1)、假设 2.2和假设 3.1-3.4 成立，则有

$$(1)\ \sum_{t=1}^T\|\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\|_F^2 = O_p\left(Tp_1p_2+Tp_1p_2^2w_c^{(s)}\right);$$

$$(2)\ \sum_{t=1}^T\|\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\|_F^2 = O_p\left(Tp_1p_2+Tp_1^2p_2w_r^{(s)}\right).$$

证明。(1) 由于 $\sum_{t=1}^T\|\mathbb{E}_t\mathbb{C}\|_F^2=\sum_{t=1}^T\sum_{i=1}^{p_1}(\sum_{k=1}^{p_2}e_{t,\ ik}c_k)^2\le c\sum_{t=1}^T\sum_{i=1}^{p_1}\sum_{k}^{p_2}{}_{,\ k'=1}e_{t,\ ik}e_{t,\ ik'}=O_p(Tp_1p_2)$，$\sum_{t=1}^T\|\mathbb{E}^t\|_F^2=\sum_{t=1}^T\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}e_{t,\ ij}^2=O_p(Tp_1p_2)$，因此

$$\sum_{t=1}^T\|\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\|_F^2\le\sum_{t=1}^T\|\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2+\sum_{t=1}^T\|\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})\|_F^2=O_p\left(Tp_1p_2+Tp_1p_2^2w_c^{(s)}\right).$$

(2) $\sum_{t=1}^T\|\mathbb{R}^\top\mathbb{E}_t\|_F^2=\sum_{t=1}^T\sum_{j=1}^{p_2}(\sum_{i=1}^{p_1}r_ie_{t,\ ij})^2\le c\sum_{t=1}^T\sum_{j=1}^{p_2}\sum_{i_1,\ i_2=1}^{p_1}e_{t,\ i_1j}e_{t,\ i_2j}=O_p(Tp_1p_2)$,

$$\sum_{t=1}^T\|\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\|_F^2\le\sum_{t=1}^T\|\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)})\|_F^2+\sum_{t=1}^T\|\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2=O_p\left(Tp_1^2p_2w_r^{(s)}+Tp_1p_2\right).$$

□

引理 S2.8。 假设 T， p_1， p_2趋于无穷，且 $m_1=k_1$， $m_2=k_2$为固定值。若假设 2.1（1）、假设 2.2以及假设 3.1-3.4 成立，则有

- (1)
$$\|\sum_{t=1}^T\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\mathbf{R}\|_F^2=O_p\left(Tp_1^2p_2^3+T^2p_1p_2^2+(Tp_1^2p_2^4+T^2p_1p_2^4)w_c^{(s)2}\right);$$
- (2)
$$\|\sum_{t=1}^T\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\mathbf{C}\|_F^2=O_p\left(Tp_1^3p_2^2+T^2p_1^2p_2+(Tp_1^4p_2^2+T^2p_1^4p_2)w_r^{(s)}w_r^{(s+1)}\right).$$

证明。（1）为简便起见，我们固定 $k_1=k_2=m_1=m_2=1$。注意到

$$\begin{aligned}\|\sum_{t=1}^T\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\widehat{\mathbf{C}}^{(s)\top}\mathbf{E}_t^\top\mathbf{R}\|_F^2\le&\|\sum_{t=1}^T\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\widehat{\mathbf{H}}_c^{(s)\top}\mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2+\|\sum_{t=1}^T\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})\widehat{\mathbf{H}}_c^{(s)\top}\mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2\\&+\|\sum_{t=1}^T\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2+\|\sum_{t=1}^T\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2.\end{aligned}$$

对于第一项,

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\widehat{\mathbf{H}}_c^{(s)\top}\mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2 &\le \|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2\|\mathbf{C}\|_F^2\|\widehat{\mathbf{H}}_c^{(s)}\|_F^4\\ &\le p_2\sum_{i'=1}^{p_1}\sum_{j'=1}^{p_2}(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}r_ic_je_{t,ij}e_{t,i'j'})^2\\ &= O_p\left(Tp_1^2p_2^3+T^2p_1p_2^2\right),\end{aligned}$$

其中

$$\begin{aligned}\mathbb{E}(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}r_ic_je_{t,ij}e_{t,i'j'})^2 &\le c\sum_{t,s=1}^{T}\sum_{i_1,i_2=1}^{p_1}\sum_{j_1,j_2=1}^{p_2}\mathrm{Cov}(e_{t,i_1j_2}e_{t,i'j'},e_{s,i_2j_2}e_{s,i'j'})+c(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\mathbb{E}e_{t,ij}e_{t,i'j'})^2\\ &= O\left(Tp_1p_2+T^2\right).\end{aligned}$$

对于第二项,

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})\widehat{\mathbf{H}}_c^{(s)\top}\mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2 &\le \|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{C}^\top\mathbf{E}_t^\top\boldsymbol{R}\|_F^2\|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2\|\widehat{\mathbf{H}}_c^{(s)}\|_F^2\\ &= O_p\left((Tp_1^2p_2^3+T^2p_1p_2^2)w_c^{(s)}\right).\end{aligned}$$

对于第三项,

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2 &\le \|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{C}\mathbf{R}^\top\mathbf{E}_t\|_F^2\|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2\|\widehat{\mathbf{H}}_c^{(s)}\|_F^2\\ &= \sum_{i'=1}^{p_1}\sum_{j'=1}^{p_2}(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}r_ic_je_{t,i'j}e_{t,ij'})^2\|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2\|\widehat{\mathbf{H}}_c^{(s)}\|_F^2\\ &= O_p\left((Tp_1^2p_2^3+T^2p_1p_2^2)w_c^{(s)}\right),\end{aligned}$$

其中

$$\begin{aligned}\mathbb{E}(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}r_ic_je_{t,i'j}e_{t,ij'})^2 &\le c\sum_{t,s=1}^{T}\sum_{i_1i_2=1}^{p_1}\sum_{j_1j_2=1}^{p_2}\mathrm{Cov}(e_{t,i\prime j_1}e_{t,i_1j\prime},e_{s,i'j_2}e_{s,i_2j\prime})+c(\sum_{t=1}^{T}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\mathbb{E}e_{t,i'j}e_{t,ij'})^2\\ &= O\left(Tp_1p_2+T^2\right).\end{aligned}$$

对于最后一项,

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2 &\le \|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2\|\sum_{t=1}^{T}\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})^\top\mathbf{E}_t^\top\mathbf{R}\|_F^2\\ &\le \|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^4\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}e_{t,ij}\|_F^2\\ &\le \|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^4\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\sum_{k=1}^{p_2}(\sum_{t=1}^{T}\sum_{i'=1}^{p_1}r_{i'}e_{t,i'k}e_{t,ij})^2\\ &= O_p\left((T^2p_1p_2^4+Tp_1^2p_2^4)w_c^{(s)2}\right),\end{aligned}$$

其中

$$\begin{aligned}\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\sum_{k=1}^{p_2}\mathbb{E}(\sum_{t=1}^{T}\sum_{i'=1}^{p_1} r_{i'}e_{t,i'k}e_{t,ij})^2 &\le c\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\sum_{k=1}^{p_2}(\sum_{t=1}^{T}\sum_{i'=1}^{p_1}\mathbb{E}e_{t,i'k}e_{t,ij})^2\\ &+ c\sum_{i=1}^{p_1}\sum_{j=1}^{p_2}\sum_{k=1}^{p_2}\sum_{t,s=1}^{T}\sum_{i_1i_2=1}^{p_1}\mathrm{Cov}(e_{t,i_1k}e_{t,ij}, e_{s,i_2k}e_{s,ij})\\ &= O_p\left(T^2p_1p_2^2 + Tp_1^2p_2^2\right).\end{aligned}$$

(2) 注意到

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s+1)}\widehat{\mathbf{R}}^{(s)\top}\mathbf{E}_t\mathbf{C}\|_F^2 &= \|\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s+1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)} + \mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)})(\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)} + \mathbf{R}\widehat{\mathbf{H}}_r^{(s)})^\top\mathbf{E}_t\mathbf{C}\|_F^2\\ &\le \|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}\widehat{\mathbf{H}}_r^{(s)\top}\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\|_F^2 + \|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}(\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)})^\top\mathbf{E}_t\mathbf{C}\|_F^2\\ &+ \|\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s+1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)})\widehat{\mathbf{H}}_r^{(s)\top}\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\|_F^2\\ &+ \|\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s+1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)})(\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)})^\top\mathbf{E}_t\mathbf{C}\|_F^2.\end{aligned}$$

对于第一项,

$$\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}\widehat{\mathbf{H}}_r^{(s)\top}\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\|_F^2 \le \|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\|_F^2\|\mathbf{R}\|_F^2\|\widehat{\mathbf{H}}_r^{(s+1)}\|_F^2\|\widehat{\mathbf{H}}_r^{(s)}\|_F^2 = O_p\left(Tp_1^3p_2^2 + T^2p_1^2p_2\right).$$

对于第二项,

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}(\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)})^\top\mathbf{E}_t\mathbf{C}\|_F^2 &\le \|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\mathbf{C}^\top\mathbf{E}_t^\top\|_F^2\|\widehat{\mathbf{R}}^{(s)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2\|\widehat{\mathbf{H}}_r^{(s+1)}\|_F^2\\ &= O_p\left((Tp_1^3p_2^2 + T^2p_1^2p_2)w_r^{(s)}\right).\end{aligned}$$

对于第三项,

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s+1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)})\widehat{\mathbf{H}}_r^{(s)\top}\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\|_F^2 &\le \|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\|_F^2\|\widehat{\mathbf{R}}^{(s+1)} - \mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}\|_F^2\|\widehat{\mathbf{H}}_r^{(s)}\|_F^2\\ &= O_p\left((Tp_1^3p_2^2 + T^2p_1^2p_2)w_r^{(s+1)}\right).\end{aligned}$$

对于最后一项,

$$
\begin{aligned}
&\|\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s+1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)})(\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)})^\top\mathbf{E}_t\mathbf{C}\|_F^2\\
&\leq\|\sum_{t=1}^{T}(\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)})^\top\mathbf{E}_t\mathbf{C}\mathbf{E}_t^\top\|_F^2\|\widehat{\mathbf{R}}^{(s+1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}\|_F^2\\
&\leq\|\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2\|\widehat{\mathbf{R}}^{(s+1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}\|_F^2\sum_{i'=1}^{p_1}\sum_{j'=1}^{p_2}\|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{C}e_{t,i'j'}\|_F^2\\
&\leq\|\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2\|\widehat{\mathbf{R}}^{(s+1)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s+1)}\|_F^2\sum_{i'=1}^{p_1}\sum_{j'=1}^{p_2}\sum_{i=1}^{p_1}(\sum_{t=1}^{T}\sum_{j=1}^{p_2}c_je_{t,ij}e_{t,i'j'})^2\\
&=O_p\left((Tp_1^4p_2^2+T^2p_1^4p_2)w_r^{(s)}w_r^{(s+1)}\right),
\end{aligned}
$$

其中

$$
\begin{aligned}
\mathbb{E}(\sum_{t=1}^{T}\sum_{j=1}^{p_2}c_je_{t,ij}e_{t,i'j'})^2&\leq c\sum_{t,s=1}^{T}\sum_{j_1,j_2=1}^{p_2}\mathrm{Cov}(e_{t,ij_1}e_{t,i'j'},e_{s,ij_2}e_{s,i'j'})+c(\sum_{t=1}^{T}\sum_{j=1}^{p_2}\mathbb{E}e_{t,ij}e_{t,i'j'})^2\\
&=O\left(Tp_2+T^2\right).
\end{aligned}
$$

□

引理 S2.9。 设 T、p_1、p_2趋于无穷大，且 $m_1=k_1$、$m_2=k_2$为固定值。若假设 2.1（1）、假设 2.2以及假设 3.1-3.4 成立，则有

$$
(1)\ \|\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\|_F^2=O_p\left(Tp_1p_2^2w_c^{(s)}+Tp_1p_2\right).
$$

$$
(2)\ \|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\|_F^2=O_p\left(Tp_1p_2^2w_c^{(s)}+Tp_1p_2\right).
$$

$$
(3)\ \|\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\mathbf{F}_t\|_F^2=O_p\left(Tp_1^2p_2w_r^{(s)}+Tp_1p_2\right).
$$

$$
(4)\ \|\sum_{t=1}^{T}\mathbf{C}^\top\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\mathbf{F}_t\|_F^2=O_p\left(Tp_1^2p_2w_r^{(s)}+Tp_1p_2\right)
$$

证明。（1）

$$
\begin{aligned}
\|\sum_{t=1}^{T}\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\|_F^2&\leq\|\sum_{t=1}^{T}\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})\mathbf{F}_t^\top\|_F^2+\|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\mathbf{F}_t^\top\|_F^2\\
&\leq\|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{F}_t^\top\|_F^2\|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2+\|\sum_{t=1}^{T}\mathbf{E}_t\mathbf{C}\mathbf{F}_t^\top\|_F^2\|\widehat{\mathbf{H}}_c^{(s)}\|_F^2\\
&=O_p\left(Tp_1p_2^2w_c^{(s)}+Tp_1p_2\right).
\end{aligned}
$$

(2)

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\widehat{\mathbf{C}}^{(s)}\mathbf{F}_t^\top\|_F^2 &\le \|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t(\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)})\mathbf{F}_t^\top\|_F^2+\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\mathbf{F}_t^\top\|_F^2\\ &\le \|\sum_{t=1}^{T}\mathbf{F}_t^\top\mathbf{R}^\top\mathbf{E}_t\|_F^2\|\widehat{\mathbf{C}}^{(s)}-\mathbf{C}\widehat{\mathbf{H}}_c^{(s)}\|_F^2+\|\sum_{t=1}^{T}\mathbf{R}^\top\mathbf{E}_t\mathbf{C}\mathbf{F}_t^\top\|_F^2\|\widehat{\mathbf{H}}_c^{(s)}\|_F^2\\ &= O_p\left(Tp_1p_2^2w_c^{(s)}+Tp_1p_2\right).\end{aligned}$$

(3)

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\mathbf{F}_t\|_F^2 &\le \|\sum_{t=1}^{T}\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)})\mathbf{F}_t\|_F^2+\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\mathbf{F}_t\|_F^2\\ &\le \|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{F}_t\|_F^2\|\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2+\|\sum_{t=1}^{T}\mathbf{E}_t^\top\mathbf{R}\mathbf{F}_t\|_F^2\|\widehat{\mathbf{H}}_r^{(s)}\|_F^2\\ &= O_p\left(Tp_1^2p_2w_r^{(s)}+Tp_1p_2\right).\end{aligned}$$

(4)

$$\begin{aligned}\|\sum_{t=1}^{T}\mathbf{C}^\top\mathbf{E}_t^\top\widehat{\mathbf{R}}^{(s)}\mathbf{F}_t\|_F^2 &\le \|\sum_{t=1}^{T}\mathbf{C}^\top\mathbf{E}_t^\top(\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)})\mathbf{F}_t\|_F^2+\|\sum_{t=1}^{T}\mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\mathbf{F}_t\|_F^2\\ &\le \|\sum_{t=1}^{T}\mathbf{C}^\top\mathbf{E}_t^\top\mathbf{F}_t\|_F^2\|\widehat{\mathbf{R}}^{(s)}-\mathbf{R}\widehat{\mathbf{H}}_r^{(s)}\|_F^2+\|\sum_{t=1}^{T}\mathbf{C}^\top\mathbf{E}_t^\top\mathbf{R}\mathbf{F}_t\|_F^2\|\widehat{\mathbf{H}}_r^{(s)}\|_F^2\\ &= O_p\left(Tp_1^2p_2w_r^{(s)}+Tp_1p_2\right).\end{aligned}$$

□