

- 你说的范数应该是矩阵的谱半径 (spectral radius) 吧, 不过对于对称正定矩阵的 L_2 范数, 二者相等。see https://en.wikipedia.org/wiki/Matrix_norm
- Lemma 1 中的 (9) 式按矩阵的谱半径处理, 相当于:

$$\max_{1 \leq i \leq n} \left\{ \frac{1}{1 - x_i} - 1 \right\} = \boxed{\max_{1 \leq i \leq n} \left\{ \frac{x_i}{1 - x_i} \right\} = \frac{\max_{1 \leq i \leq n} \{x_i\}}{1 - \max_{1 \leq i \leq n} \{x_i\}}} \quad (1)$$

所以你的 Lemma 1 可能还是有问题。

- 关于我的仿真, 规模比较小, 具体参数并不重要, 主要想说明如果层数考虑的太多, 由于指数衰减的特性可能最后面几层会由于数值误差的影响结果有问题。当然这一点我并没有具体验证。