Factorized Machine (FM)

王树森

线性模型

- 有 d 个特征，记作 $\mathbf{x} = [x_1, \cdots, x_d]$。
- 线性模型：

$$\underline{p} = \underline{b + \sum_{i=1}^{d} w_i x_i}.$$

- 模型有 $d + 1$ 个参数：$\mathbf{w} = [w_1, \cdots, w_d]$ 和 b。
- 预测是特征的加权和。（只有加，没有乘。）

二阶交叉特征

- 有 d 个特征，记作 $\mathbf{x} = [x_1, \cdots, x_d]$。
- 线性模型 + 二阶交叉特征：

$$p = \underline{b + \sum_{i=1}^{d} w_i x_i} + \sum_{i=1}^{d} \sum_{j=i+1}^{d} \underline{u_{ij}} x_i x_j .$$

- 模型有 $O(d^2)$ 个参数。

二阶交叉特征

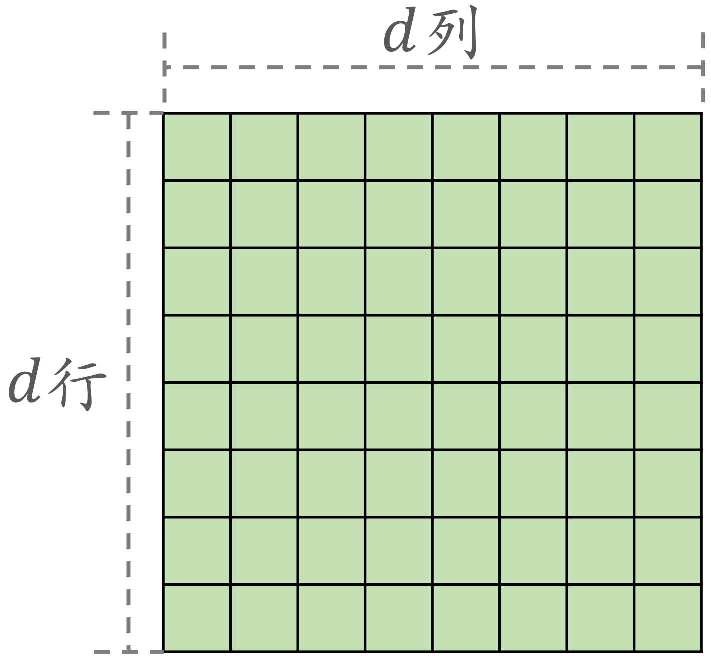
- 线性模型 + 二阶交叉特征：

$$p = b + \sum_{i=1}^{d} w_i x_i + \sum_{i=1}^{d} \sum_{j=i+1}^{d} u_{ij} x_i x_j.$$

二阶交叉特征

- 线性模型 + 二阶交叉特征：

$$p = b + \sum_{i=1}^{d} w_i x_i + \sum_{i=1}^{d} \sum_{j=i+1}^{d} u_{ij} x_i x_j .$$

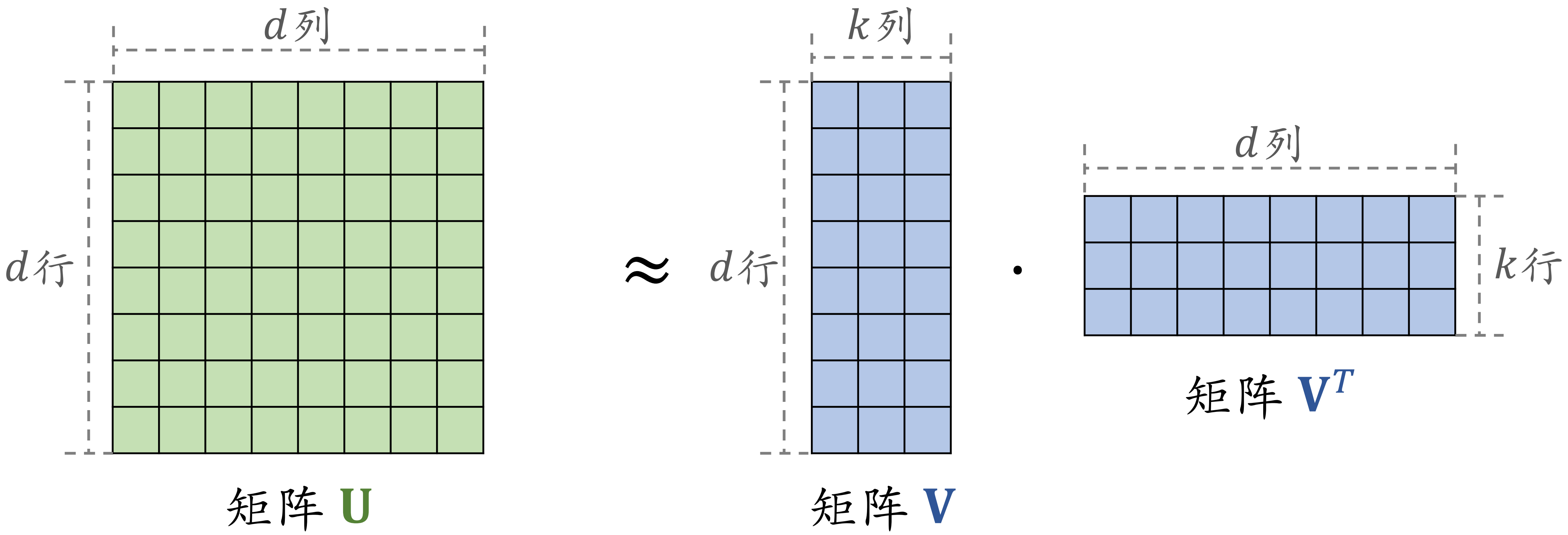


矩阵 **U**

二阶交叉特征

- 线性模型 + 二阶交叉特征：

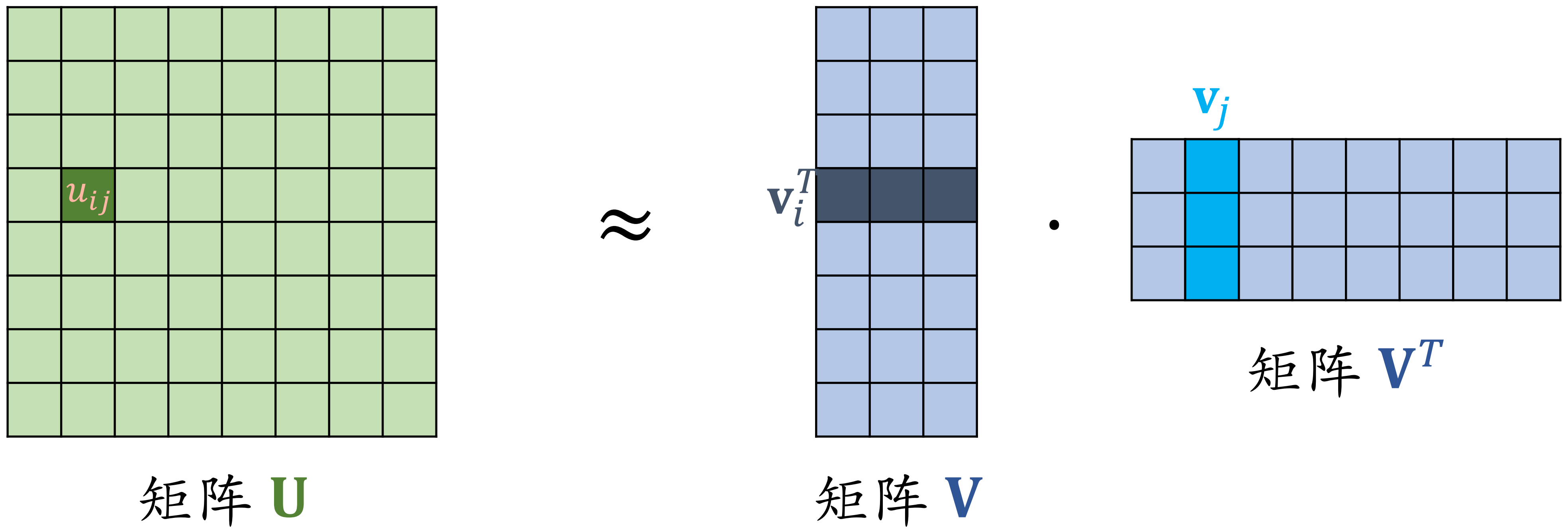
$$p = b + \sum_{i=1}^{d} w_i x_i + \sum_{i=1}^{d} \sum_{j=i+1}^{d} u_{ij} x_i x_j.$$



二阶交叉特征

- 线性模型 + 二阶交叉特征：

$$p = b + \sum_{i=1}^{d} w_i x_i + \sum_{i=1}^{d} \sum_{j=i+1}^{d} \underbrace{u_{ij}}_{\approx \mathbf{v}_i^T \mathbf{v}_j} x_i x_j.$$



二阶交叉特征

- 线性模型 + 二阶交叉特征：

$$p = b + \sum_{i=1}^{d} w_i x_i + \sum_{i=1}^{d} \sum_{j=i+1}^{d} \underbrace{u_{ij}}_{\approx \mathbf{v}_i^T \mathbf{v}_j} x_i x_j.$$

- Factorized Machine (FM)：

$$p = b + \sum_{i=1}^{d} w_i x_i + \sum_{i=1}^{d} \sum_{j=i+1}^{d} \underline{(\mathbf{v}_i^T \mathbf{v}_j)} x_i x_j.$$

- FM 模型有 $O(kd)$ 个参数。（$k \ll d$）

Factorized Machine

- FM 是线性模型的替代品，能用线性回归、逻辑回归的场景，都可以用 FM。
- FM 使用二阶交叉特征，表达能力比线性模型更强。

Factorized Machine

- FM 是线性模型的替代品，能用线性回归、逻辑回归的场景，都可以用 FM。
- FM 使用二阶交叉特征，表达能力比线性模型更强。
- 通过做近似 $u_{ij} \approx \mathbf{v}_i^T \mathbf{v}_j$，FM 把二阶交叉权重的数量从 $O(d^2)$ 降低到 $O(kd)$。

参考文献：

- Steffen Rendle. Factorization machines. In *ICDM*, 2010.

Thank You!