G 一个 $n \times m$ 矩阵，n是样本量，m是SNP标记数量

M 一个 $n \times m$ 矩阵，其每一列的原素都是相同的。

$$M = \begin{bmatrix} 2p_1 & 2p_2 & \cdots & 2p_m \\ 2p_1 & 2p_2 & \cdots & 2p_m \\ \vdots & \vdots & & \vdots \\ 2p_1 & 2p_2 & \cdots & 2p_m \end{bmatrix}$$

注意，M的每一行是相同的

V 一个 scalar $V = \sum_{i=1}^{m} 2p_i q_i$

$$GRM = \frac{1}{V}(G-M)(G-M)^T \quad (Eq1).$$

GRM是一个 $n \times n$ 矩阵。

$$\begin{bmatrix} GRM \end{bmatrix}_{N \times N}$$

将Eq1展开。

$$GRM = \frac{1}{V}(G \cdot G^T - M \cdot G^T - G M^T + M \cdot M).$$

为表达方便，舍去 $\frac{1}{V}$.

$$GRM = G \cdot G^T - M \cdot G^T - G \cdot M^T + M \cdot M$$

$M \cdot G^T$ 是 $G \cdot M^T$ 的转秩，所以只要计算一个就可以。

如上提到，M每一行相同。

$M \cdot G^T$ 的计算和存储都可以压缩。

只要计算M第一行与 G^T 的乘积，且存为一个Vector。

$M_{[1,]} \cdot G^T$ 包含所有 $M \cdot G^T$ ($G \cdot M^T$) 信息。

$M \cdot M$ 可以压缩为一个元素 $\sum_{i=1}^{m} (2p_i)^2$

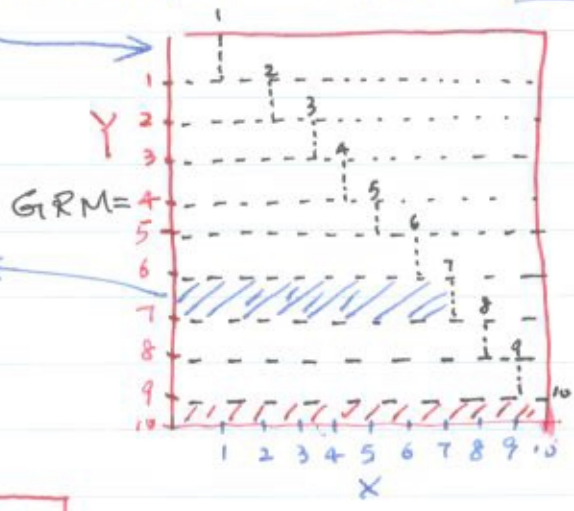
综上，我们只要关心 $G \cdot G^T$ 相乘作。

$G \cdot G^T$ 的 mailman 应用.（第一种分解）

G 矩阵的所有元素是 {0, 1, 2}. 元素数量 S=3

G 是 n·m 矩阵. 假定 n=100, $m=3^{10}=59049$.

GRM=$G \cdot G^T$ 的 100×100 矩阵可分解成 100 个单元格.（方块）

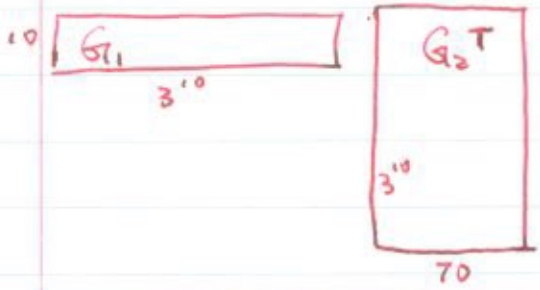


从阴影这条为例

是 $G_1 \cdot G_2^T$

G_1 是 $10 \cdot 3^{10}$ 矩阵.

G_2^T 是 $3^{10} \cdot 70$ 矩阵.



这个矩阵相乘可以用 mailman 操作.

将 G_1 分解成 U·P, 与 G_2^T 每列相乘

同理. GRM 最后一条则是（红色阴影）

$G_1 = 10 \cdot 3^{10}$

$G_2^T = 3^{10} \cdot 100$

缺点：随着 Y 轴从 1 向 10 延伸. mailman 计算优势越大. 因为 G_1 的分解 O(m·n) 分摊到最多 n 个 vector 中.

从更加细微角度入手. 当 $G_1 \cdot G_1^T$ 时. 计算量可以减半.
（GRM 最后格那个方块）

$G_1 \cdot G_1^T$ 的第二种分解.

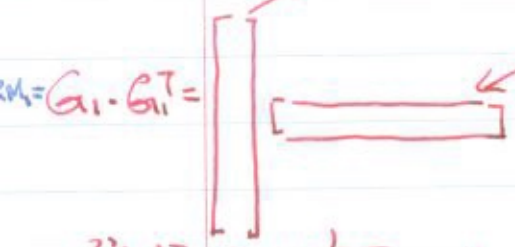
注意. G矩阵 行列到的元素是相同的。都是 {1, 2, 3}. 所以可以有相对标记的解析.

假定. $n = 3^8 = 6561$. $m = 80$.

则将 G 垂直水平拆成 10 等份.

$GRM_1 = G_1 \cdot G_1^T$

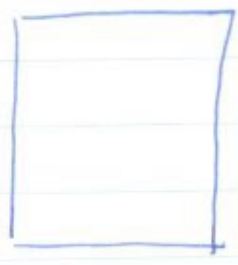
操作见左侧示图



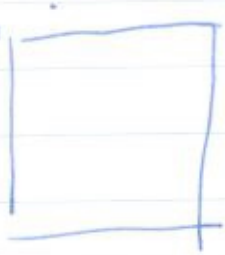
依旧可以套用 mailman

但要先转秩. 见朱志翔对 mailman 说明书.

$GRM_2 = G_2 \cdot G_2^T$



$GRM_{10} = G_{10} \cdot G_{10}^T$



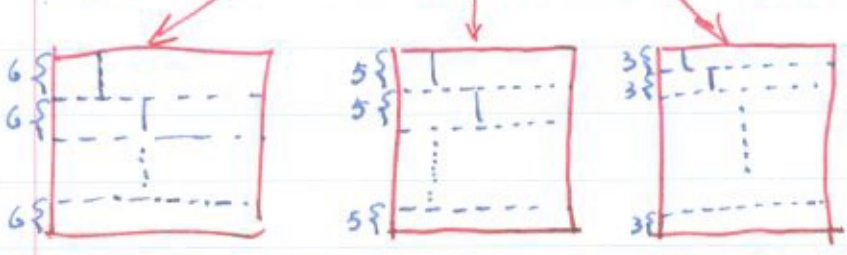
$$GRM = \sum_{K=1}^{10} GRM_K$$

综合考虑，对于 GRM 的计算 有三种 可能的方式

方式 1) 对 GRM 进行多格化分割.
也就是对 G 的 m 个 SNP 标记从 S^k 进行分割.

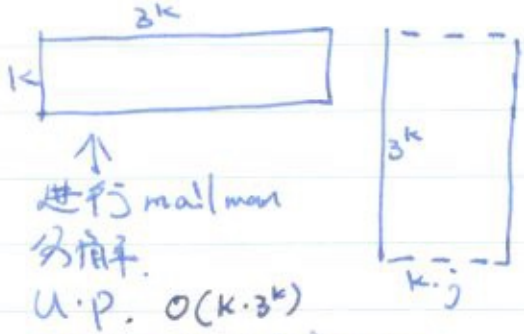
∵ $m = 1000$
$= 3^6 + 3^5 + 3^3 + 1$
$= 729 + 243 + 27 + 1$

∴ $GRM = GRM_6 + GRM_5 + GRM_3 + GRM_1$.



① 注意，对于 GRM_k，最后一个分割 可能需要增加一行或几行（最多 k-1 行）。

② 对于 GRM_k，第 j 个横向分割 可以写为
一个 $k \cdot 3^k$ 矩阵 与 $3^k \cdot (k \cdot j)$ 矩阵相乘.



进行 mailman 分解.
U.P. $O(k \cdot 3^k)$
然后与右侧相乘.

a) 共 SNP 无论 j 取值，左侧矩阵总是 $k \cdot 3^k$，所以
总的 mailman 分解 需要 $O(k \cdot 3^k \cdot \lceil \frac{n}{k} \rceil) \approx O(n \cdot 3^k)$.

b) 针对不同 k 取值，总共的 mailman 分解
$O(\sum_k n \cdot 3^k) = O(n \cdot m)$，是 constant，与 SNP 分组数无关。