Homework 4

Ido Nutov & Tuvy Lemberg

July 12, 2025

1

Let $G \leq S_n$ be a subgroup of the permutation group S_n acting on the index set $X = \{1, 2, \ldots, n\}$ where n is a finite natural number. An element g acts on a vector $v \in \mathbb{R}^n$ on the coordinate i by $(gv)_i = v_{g^{-1}(i)}$. Consider a G-linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$. Thus, $T(gv) = gT(v)$. Therefore, T must fulfill $Tgv = gTv$ for all $g \in G$ and $v \in \mathbb{R}^n$. As a result $Tg = gT$ and equivalently, $g^{-1}Tg = T$ for all $g \in G$ (in the lectures we used W). Applying g from the right on the columns of T and g^{-1} from the left on the rows of T results in the following condition:

$$T_{i,j} = T_{g^{-1}(i),g^{-1}(j)}. \tag{1}$$

g acts on T and on each orbit of T the requirement is that T is constant. Thus, the condition requires that $T_{i,j} = T_{l,m}$ if $Orb_G\left((i,j)\right) = Orb_G\left((l,m)\right)$. Therefore the number of free parameters is the number of orbits.

We define $U = X \times X$. Burnside's Lemma states that

$$|U/G| = \frac{1}{|G|}\sum_{g\in G}|U^g| = \frac{1}{|G|}\sum_{g\in G}|X^g|^2 = \frac{1}{|G|}\sum_{g\in G}\chi_\rho(g)^2, \tag{2}$$

where $|U/G|$ is the number of orbits and $U^g = \{u \in U \mid gu = u\}$ is the set of elements in U that g doesn't change. We have the property of $|U^g| = \chi_\sigma(g)$,

2 Question 2

2.1

The natural group here is the cross product of the permutation groups $S_n \times S_m$ acts on the Netflix rating matrices $A \in \mathbb{R}^{n\times m}$. It can be explained by the fact that the order of the m movies and the order of the n users is arbitrary and both can be changed independently.

2.2

Since we look at the G-linear transformation from matrices in $\mathbb{R}^{n\times m}$ to itself, it will be more convenient to consider the vectorization of these matrices. Thus

$\text{vec}(A) = [a_{1,1}, \dots, a_{n,1}, \dots, a_{1,m}, \dots, a_{n,m}]^T \in \mathbb{R}^{nm}$. Let $g = (g_l, g_r) \in S_n \times S_m$ be a group element where $g_l \in S_n$ is a left action that acts on the rows of A and $g_r \in S_m$ acts on the columns of A as follows:

$$gA = g_l A g_r^T = P_l A P_r^T, \tag{3}$$

where $P_l \in \mathbb{R}^{n\times n}$ and $P_r \in \mathbb{R}^{m\times m}$ are the permutations of g_l and g_r, respectively. Define $v = \text{vec}(A) \in \mathbb{R}^{nm}$. We are actually looking on the isomorphic group $G \leq S_{nm}$, which acts on the indexes of the vectorization of A. We will use G and $S_n \times S_m$ interchangeably. In index notation we have $(gA)_{i,j} = A_{g_l^{-1}(i), g_r^{-1}(j)}$. We are looking at the affine transformation that is equivariant, thus $f(v) = Wv + b$, that fulfills for all $g \in S_n \times S_m$ and $v \in \mathbb{R}^{nm}$

$$\begin{aligned} f(gv) &= W(gv) + b = g(Wv + b) \\ &\Rightarrow Wgv + b = gWv + gb \end{aligned} \tag{4}$$

Since (4) needs to be fulfilled for all $v \in \mathbb{R}^{nm}$ it leads to $Wg = gW$ and equivalently $gWg^T = W$. Moreover, it requires that $gb = b$ for all $g \in G$. We have

$$gA = g_l A g_r^T = P_l A P_r^T. \tag{5}$$

By vectorizing the above we obtain

$$\text{vec}\left(P_l A P_r^T\right) = (P_r \otimes P_l)\,\text{vec}(A) \Rightarrow P_{nm} = (P_r \otimes P_l) \tag{6}$$

For b, we can see from (4) that this means that b is constant, $b = c \cdot \mathbf{1}$. This can be seen by applying the inverse vectorization operation on b. Since the condition (4) requires that b will be invariant to the action of g, and by acting on b we can replace each index $\text{vec}^{-1}(b_{i,j})$ with any other index. Therefore, all values of b required to be the same.

Let us look at the action of the permutation $g^{-1} = (g_l^{-1}, g_r^{-1})$ on the vector v. We take the inverse g^{-1} for convenience of notation. Originally, the element with the index $A_{i,j}$ moves to $v_{i+n(j-1)}$, therefore

$$\left(g^{-1}A\right)_{i,j} = A_{g_l(i), g_r(j)} = v_{g_l(i)+n(g_r(j)-1)}. \tag{7}$$

Consequently, for an index $1 \leq k \leq nm$, define $j = k//n$ as the truncating integer division and $i = k\%n$ as the modulo operation, thus $\left(g^{-1}v\right)_k = v_{g_l(i)+n\cdot(g_r(j)-1)}$. Let k, l be indexes in $[1, \dots, nm]$ with the decomposition $k = i_1 + n(j_1 - 1)$ and $l = i_2 + n(j_2 - 1)$. Then $\left(g^{-1}Wg\right)_{k,l} = W_{g_l(i_1)+n\cdot(g_r(j_1)-1), g_l(i_2)+n(g_r(j_2)-1)}$.

2.2.1 Orbit analysis

Let us look at the possible orbits in the index set $X^2 = \{1, 2, \dots, nm\}^2$. Indexes (k', l') can be on the same orbits as (k, l) if and only if there exists a permutation $(g_l, g_r) \in S_n \times S_m$ such that

$$\begin{cases} k' &= g_l(i_1) + n\left(g_r(j_1) - 1\right) \\ l' &= g_l(i_2) + n\left(g_r(j_2) - 1\right) \end{cases}. \tag{8}$$

Define i_3, i_4, j_3, j_4 such that $k' = i_3 + n(j_3 - 1)$ and $l' = i_4 + n(j_4 - 1)$. It is clear that we can create equivalence groups by the rule:

$$\left(i_1 \overset{\neq}{=} i_2\right) \wedge \left(j_1 \overset{\neq}{=} j_2\right). \tag{9}$$

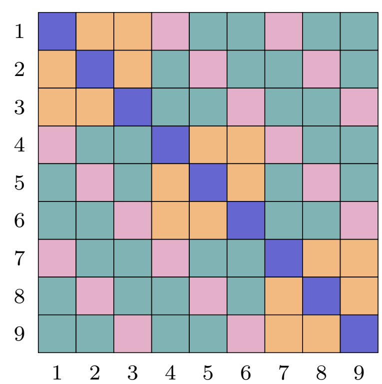
It is clear that $i_3 = i_4$ holds if and only if $i_1 = i_2$, since by applying g_l we have $g_l(i_1) = g_l(i_2) = i_3 = i_4$, and since g_l is a bijection. The same goes for $j_3 = j_4$.

These are the orbits and the explanation for why they are the only orbits is as follows: if $i_1 = i_2$ then we can move to any other k', l' that has $i_3 = i_4$ by applying the right $g_l(i_1) = i_3$. If $i_1 \neq i_2$ we can move to any other indexes $i_3 \neq i_4$ since S_n is 2-transitive. The same goes for the indexes j. Therefore, we will have only these 4 orbits:

$$\begin{aligned}
O_{=i=j} &: i_1 = i_2, j_1 = j_2 \\
O_{=i\neq j} &: i_1 = i_2, j_1 \neq j_2 \\
O_{\neq i=j} &: i_1 \neq i_2, j_1 = j_2 \\
O_{\neq i\neq j} &: i_1 \neq i_2, j_1 \neq j_2
\end{aligned}$$

2.3

For $n = m = 3$ we will obtain $W \in \mathbb{R}^{9\times 9}$ and $b \in \mathbb{R}^9 = c \cdot \mathbf{1}$



2.4

In the case of a linear invariant layer, for all $v \in \mathbb{R}^{nm}$ and $g \in G$ we have:

$$\begin{aligned}
f(gv) &= f(v) \\
\Rightarrow w^T gv + b &= w^T v + b \\
\Rightarrow g^T w &= w
\end{aligned}$$

For $w \in \mathbb{R}^{nm}$, this condition is equivalent to the condition on b in the previous clause. Therefore, $w = c_2 \cdot \mathbf{1}$.

2.5

In the case of 3-tensors with the symmetric group is $S_{N_1} \times S_{N_2} \times S_{N_3}$ applied independently to each index axis of the tensor. We have:

$$v = \text{vec}(A) \;\; v_k = a_{i_1,i_2,i_3}, \tag{10}$$

where $k = i_1+N_1(i_2-1)+N_1N_2(i_3-1)$. Define $k = i_1^k+N_1(i_2^k-1)+N_1N_2(i_3^k-1)$ and $l = i_1^l + N_1(i_2^l - 1) + N_1N_2(i_3^l - 1)$. We will have the same derivation as in the previous clause 2.2, thus

$$W_{k,l} = W_{g(k),g(l)}$$

The indexes (k, l) can transform to

$$\begin{aligned} k' &= g_1(i_1^k) + N_1(g_2(i_2^k) - 1) + N_1N_2(g_3(i_3^k) - 1) \\ l' &= g_1(i_1^l) + N_1(g_2(i_2^l) - 1) + N_1N_2(g_3(i_3^l) - 1) \end{aligned}$$

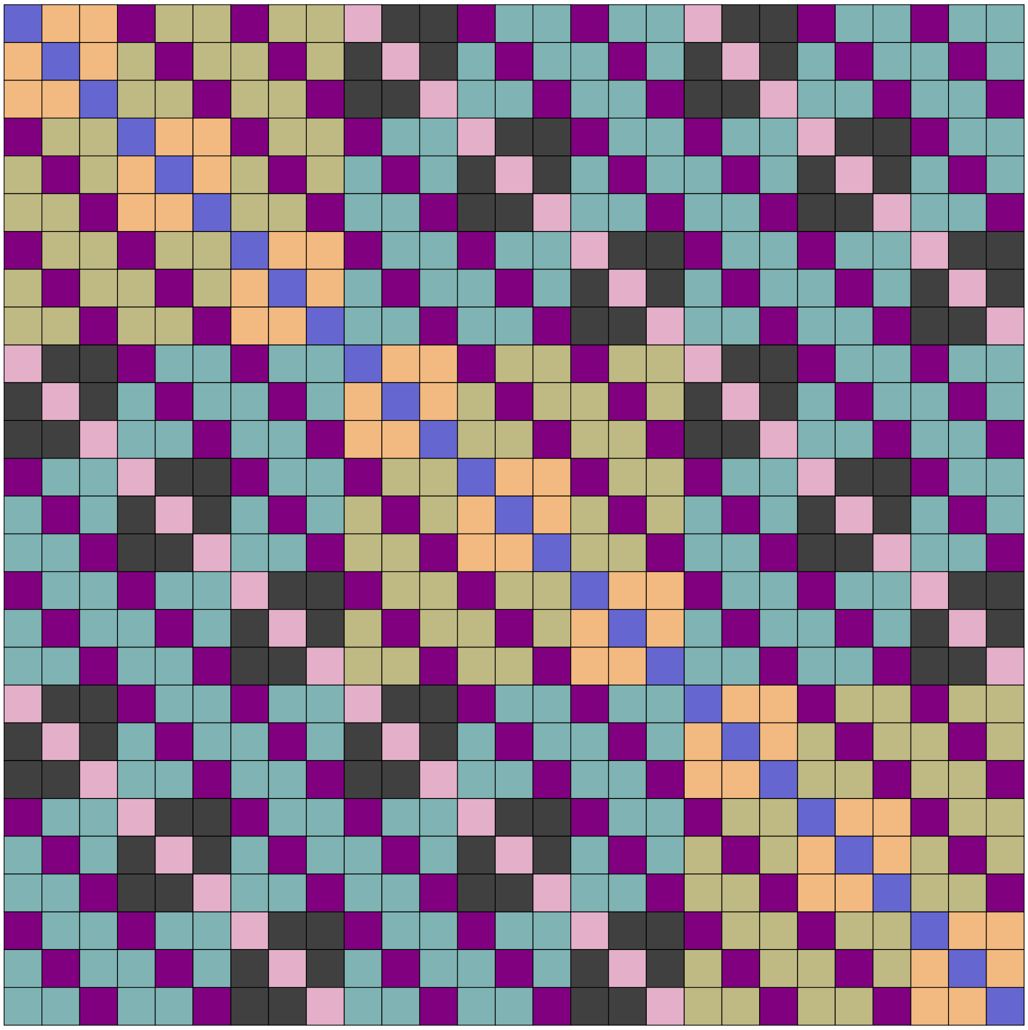
where the equivalence relations follow the same pattern as before, thus the orbits are

$$\bigwedge_{n=1}^{3} i_n^k \overset{\neq}{=} i_n^l. \tag{11}$$

As a result, the orbits are:

$$\begin{aligned} O_{=i_1=i_2=i_3} &: i_1^k = i_1^l, i_2^k = i_2^l, i_3^k = i_3^l \\ O_{=i_1=i_2\neq i_3} &: i_1^k = i_1^l, i_2^k = i_2^l, i_3^k \neq i_3^l \\ O_{=i_1\neq i_2=i_3} &: i_1^k = i_1^l, i_2^k \neq i_2^l, i_3^k = i_3^l \\ &\vdots \\ O_{\neq i_1\neq i_2\neq i_3} &: i_1^k \neq i_1^l, i_2^k \neq i_2^l, i_3^k \neq i_3^l \end{aligned}$$

For $N_1 = N_2 = N_3 = 3$, we obtain $W \in \mathbb{R}^{27\times 27}$ and $b \in \mathbb{R}^{27}$.



3

4