involve
a journal of mathematics

The monochromatic column problem with a prime number of colors

Loran Crowell and Steve Szabo

msp

2019 vol. 12, no. 8

INVOLVE 12:8 (2019)
dx.doi.org/10.2140/involve.2019.12.1415

The monochromatic column problem with a prime number of colors

Loran Crowell and Steve Szabo

(Communicated by Kenneth S. Berenhaut)

Let $p_1, \dots, p_n$ be a sequence of n pairwise coprime positive integers, $P = p_1 \cdots p_n$, and $0, \dots, m-1$ be a sequence of m different colors. Let A be an $n \times mP$ matrix of colors in which row i consists of blocks of p_i consecutive entries of the same color with colors 0 through $m-1$ repeated cyclically. The monochromatic column problem is to determine the number of columns of A in which every entry is the same color. The solution for a prime number of colors is provided.

1. Introduction

Let m be a positive integer. The colors for m are represented by the integers $0, 1, \dots, m-1$. An $n \times s$ *m-color matrix* is an $n \times s$ matrix $A = (a_{ij})$ in which every entry is one of the m colors. Column j of A is *monochromatic* if $a_{ij} = a_{1j}$ for $1 \le i \le n$. For a positive integer p, row i of A is p-blocked with initial color ρ if $p \mid s$ and, for $1 \le j \le s$,

$$a_{ij} = \left(\left\lfloor \frac{j-1}{p} + \rho \right\rfloor \right) \bmod m.$$

For $D = \{(p_i, \rho_i)\}_{i=1}^{n}$, where $p_1, \dots, p_n$ are pairwise coprime positive integers and $\rho_i \in \{0, \dots, m-1\}$, an $n \times mp_1 \cdots p_n$ m-color matrix A is the *(m, D)-color matrix* if for every i satisfying $1 \le i \le n$, row i of A is p_i-blocked with initial color ρ_i. For instance, the layout of the $(5, \{(2, 1), (3, 4)\})$-color matrix is

$$\begin{pmatrix} 1&1&2&2&3&3&4&4&0&0&1&1&2&2&3&3&4&4&0&0&1&1&2&2&3&3&4&4&0&0 \\ 4&4&4&0&0&0&1&1&1&2&2&2&3&3&3&4&4&4&0&0&0&1&1&1&2&2&2&3&3&3 \end{pmatrix}.$$

MSC2010: 05A15, 11A07.
Keywords: monochromatic column problem, Chinese remainder theorem, multiple sequence alignment problem.

The *monochromatic column problem* (*MCP*) is to determine the number of monochromatic columns in the (m, D)-color matrix, which is denoted by $N(m, D)$. Note, $N(5, \{(2, 1), (3, 4)\}) = 6$.

The MCP was originally posed in [Nagpaul and Jain 2002]. Their stated motivation is captured in the following from their paper:

> The motivation for studying this problem arose from a question asked by a biomathematician working on the multiple sequence problem that deals with finding, for given k sequences of characters from a fixed alphabet, an alignment with optimal score according to a given scoring scheme.

The multiple sequence alignment problem is a well-studied problem in molecular biology. It is of crucial importance according to [Jiang et al. 1999]. Independently of its tenuous connections to the multiple sequence alignment problem, the MCP is an interesting combinatorial problem in its own right.

The solution of the MCP for two colors is given in [Nagpaul and Jain 2002] and for three colors is given in [Srivastava and Szabo 2008]. The technique developed by Srivastava and Szabo for three colors is generalized here to give the solution for a prime number colors. A partial solution for the prime color case was the topic of [Crowell 2016].

Section 2 contains the complete solution to the prime color problem. In Section 3, the three color solution is restated, correcting a small issue in the solution in [Srivastava and Szabo 2008].

2. The monochromatic column problem: a prime number of colors

Throughout this section, let n be a positive integer, q a prime, $D = \{(p_i, \rho_i)\}_{i=1}^{n}$, where $p_1, \dots, p_n$ are pairwise coprime positive integers and $\rho_i \in \{0, \dots, q-1\}$, and $A = (a_{ij})$ be the (q, D)-color matrix. To solve the prime color problem, three cases are considered which exhaust the possibilities. First, in Proposition 1, it is assumed that $p_1, \dots, p_n$ are congruent to one another modulo q. Then in Proposition 2, it is assumed that $p_1, \dots, p_n$ may not all be congruent to one another but none are divisible by q. Finally, in Proposition 3, it is assumed that $q \mid p_n$. In the statements of the propositions, an ordering of $p_1, \dots, p_n$ is assumed, but of course this ordering does not affect the number of monochromatic columns.

Proposition 1. *Let $s \in \{1, \dots, q-1\}$. Assume $p_i \equiv s \pmod q$ for $i \in \{1, \dots, n\}$. Then*

$$N(q, D) = q \sum_{\beta=1}^{\min\{p_1, q\}} \prod_{i=1}^{n} \frac{p_i - s}{q} + \left\lfloor \frac{s - (\beta + s(\rho_i - \rho_1) - 1) \bmod q - 1}{q} \right\rfloor + 1.$$

Proof. Let $P = p_1 p_2 \cdots p_n$. Note

$$a_{ij} = \left(\left\lfloor \frac{j-1}{p_i} \right\rfloor + \rho_i \right) \bmod q$$

since $\lfloor (j-1)/p_i \rfloor$ calculates the number of complete blocks that the element is away from the beginning of the row. This number of blocks is added to the starting color of the row, ρ_i, and then taking this modulo q gives the color. We will first show that if column j is monochromatic, then a column some multiple of P away is also monochromatic.

Let $1 \le x, y \le n,\ 1 \le j \le P$, and $0 \le \alpha \le q-1$. Since $p_x \equiv p_y \pmod{q}$,

$$\frac{\alpha P}{p_x} \equiv \frac{\alpha P}{p_y} \pmod{q}$$

$$\begin{aligned} \left\lfloor \frac{j-1}{p_x} \right\rfloor - \frac{\alpha P}{p_x} - \left\lfloor \frac{j-1}{p_x} \right\rfloor &\equiv \left\lfloor \frac{j-1}{p_y} \right\rfloor - \frac{\alpha P}{p_y} - \left\lfloor \frac{j-1}{p_y} \right\rfloor \pmod{q} \\ \left\lfloor \frac{j-1}{p_x} \right\rfloor - \left\lfloor \frac{\alpha P + j - 1}{p_x} \right\rfloor &\equiv \left\lfloor \frac{j-1}{p_y} \right\rfloor - \left\lfloor \frac{\alpha P + j - 1}{p_y} \right\rfloor \pmod{q} \end{aligned} \tag{1}$$

$$a_{x,j} - a_{x,\alpha P + j} \equiv a_{y,j} - a_{y,\alpha P + j} \pmod{q}.$$

This shows that if column j is monochromatic, then so is column $\alpha P + j$. Hence, it suffices to count the number of monochromatic columns in the first P columns of A and multiply by q.

Let

$$k_{ij} = j - \left\lfloor \frac{j-1}{p_i} \right\rfloor p_i.$$

This is the count into the $\lfloor (j-1)/p_i \rfloor$-th monocolored block in the i-th row.

Since $p_1, \dots, p_n$ are pairwise coprime integers and $1 \le k_{ij} \le p_i$, the Chinese remainder theorem guarantees that $|\{(k_{1j}, \dots, k_{nj})\}_{j=1}^{P}| = P$. Therefore, by counting the n-tuples that map to a monochromatic column, the number of monochromatic columns in the first P columns of A can be determined. For $1 \le i \le n$ and $1 \le j \le P$,

$$a_{ij} = \left(\left\lfloor \frac{j-1}{p_i} \right\rfloor + \rho_i \right) \bmod q = \left(\frac{j - k_{ij}}{p_i} + \rho_i \right) \bmod q.$$

Since $p_i \equiv s \pmod{q}$, we have $a_{ij} = a_{1j}$ if and only if $k_{ij} \equiv k_{1j} + s(\rho_i - \rho_1) \pmod{q}$. So, column j is monochromatic if and only if $k_{ij} \equiv k_{1j} + \rho_i s \pmod{q}$ for all $i \in \{1, \dots, n\}$. Hence, the number of monochromatic columns in the first P columns of A is the product of the number of integer solutions to

$$1 \le q x_i + (k_{1j} + s(\rho_i - \rho_1)) \bmod q \le p_i$$

for each $i \in \{1, \ldots, n\}$; equivalently,

$$\frac{1-(k_{1j}+s(\rho_i-\rho_1)) \bmod q}{q} \le x_i \le \frac{p_i-(k_{1j}+s(\rho_i-\rho_1)) \bmod q}{q}.$$

The number of integer solutions for a given i is

$$\begin{aligned}
&\left\lfloor \frac{p_i-(k_{1j}+s(\rho_i-\rho_1)) \bmod q}{q} \right\rfloor - \left\lceil \frac{1-(k_{1j}+s(\rho_i-\rho_1)) \bmod q}{q} \right\rceil + 1 \\
&= \frac{p_i-s}{q} + \left\lfloor \frac{s-(k_{1j}+s(\rho_i-\rho_1)) \bmod q}{q} \right\rfloor + \left\lfloor \frac{(k_{1j}+s(\rho_i-\rho_1)) \bmod q-1}{q} \right\rfloor + 1 \\
&= \frac{p_i-s}{q} + \left\lfloor \frac{s-(k_{1j}+s(\rho_i-\rho_1)-1) \bmod q-1}{q} \right\rfloor + 1.
\end{aligned}$$

The possible values of $[(k_{1j}+s(\rho_i-\rho_1))-1] \bmod q$ are given by

$$\{(\beta+s(\rho_i-\rho_1)-1) \bmod q \mid \beta \in \{1, \ldots, \min\{p_1, q\}\}.$$

Summing over these possibilities for k_{1j}, multiplying the number of solutions for each row, and multiplying the sum by q, we find that the number of monochromatic columns in A is

$$N(q, D) = q \sum_{\beta=1}^{\min\{p_1,q\}} \prod_{i=1}^{n} \frac{p_i-s}{q} + \left\lfloor \frac{s-(\beta+s(\rho_i-\rho_1)-1) \bmod q-1}{q} \right\rfloor + 1. \quad \square$$

Proposition 2. *Let* $S = \{p_i \bmod q \mid i \in I\}$, $r = |S|$ *and* $s_1, \ldots, s_r \in S$ *be the distinct elements of* S. *Assume* $q \nmid p_i$ *for* $i \in I$ *and* $r > 1$. *Let* $i_0, i_1, \ldots, i_r$ *be such that* $i_0 = 0$, $i_r = n$, *and* $p_i \equiv s_l$ *for* $i_{l-1} < i \le i_l$. *Let*

$$B = \{(\beta_1, \ldots, \beta_r) \mid \beta_l \in \{1, \ldots, \min\{p_{i_l}, q\}\}\}, \tag{2}$$

where

$$\beta_l = \frac{\beta_1(s_l-s_2)+\beta_2(s_l-s_1)+s_l(s_2-s_1)(\rho_{i_l}-\rho_{i_1})+s_2(s_l-s_1)(\rho_{i_1}-\rho_{i_2})}{s_2-s_1}.$$

Then

$$N(q, D) = \sum_{(\beta_1,\ldots,\beta_r)\in B} \prod_{l=1}^{r} \prod_{i=i_{l-1}+1}^{i_l} \frac{p_i-s_l}{q} + \left\lfloor \frac{s_l-(\beta_l+s_l(\rho_i-\rho_{i_l})-1) \bmod q-1}{q} \right\rfloor + 1.$$

Proof. Let $P = p_1 \cdots p_n$. From the proof of Proposition 1, the following can be deduced. The number of columns in the first P columns of A such that the color vector of the column, $(c_1, \ldots, c_n)$, has the property that $c_i = c_{i_l}$ for $i_{l-1} < i \le i_l$ (i.e., is a column where the colors are identical if the associated p_i's are congruent) is

$$\sum_{(\beta_1,\ldots,\beta_r)\in B'} \prod_{l=1}^{r} \prod_{i=i_{l-1}+1}^{i_l} \frac{p_i - s_l}{q} + \left\lfloor \frac{s_l - (\beta_l + s_l(\rho_i - \rho_{i_l}) - 1) \bmod q - 1}{q} \right\rfloor + 1,$$

where

$$B' = \{(\beta_1, \ldots, \beta_r) \mid \beta_l \in \{1, \ldots, \min\{p_{i_l}, q\}\}\}.$$

Such columns will be called r-chromatic columns. First, it is shown that for $1 \le \alpha \le q-1$ and $1 \le j \le P$, column j is r-chromatic if and only if column $j+\alpha P$ is r-chromatic. Fix l and let $i_l - 1 \le x, y \le i_l$, $1 \le j \le P$, and $0 \le \alpha \le q-1$. Since $p_x \equiv p_y \pmod q$, the computations of (1) hold and we have

$$a_{x,j} - a_{x,\alpha P+j} \equiv a_{y,j} - a_{y,\alpha P+j} \pmod q.$$

This shows that column j is r-chromatic if and only if column $\alpha P + j$ is r-chromatic. Next, the conditions on an r-chromatic column, j, that guarantee that one and only one of the set of columns $\{j, j+P, \ldots, (q-1)P\}$ is monochromatic is developed. Let $j \in \{1, \ldots, P\}$ and assume column j is r-chromatic. Denote by the r-tuple $(c_1, \ldots, c_r)$ the entries of an r-chromatic column where $a_{ij} = c_l$ for all $i \in \{i_1, \ldots, i_l\}$. Of the noted columns, the only ones that may be monochromatic will have the property that

$$c_1 + \frac{\alpha P}{s_1} \equiv c_2 + \frac{\alpha P}{s_2} \pmod q$$

for some $\alpha \in \{0, \ldots, q-1\}$. So,

$$\alpha = (c_2 - c_1)\left(\frac{P}{s_1} - \frac{P}{s_2}\right)^{q-2} \bmod q.$$

This then shows that the only possible column that may be monochromatic is $\alpha P + j$. Furthermore, for such a column to be monochromatic, working over $\mathbb{Z}_p$, for $l \in \{3, \ldots, q\}$,

$$\frac{c_l - c_1}{1/s_1 - 1/s_l} = \frac{c_2 - c_1}{1/s_1 - 1/s_2}$$

$$\frac{s_1 s_l}{s_l - s_1}(c_l - c_1) = \frac{s_1 s_2}{s_2 - s_1}(c_2 - c_1)$$

$$\left(\frac{j - k_{i_l j}}{s_l} + \rho_{i_l} - \left(\frac{j - k_{i_1 j}}{s_1} + \rho_{i_1}\right)\right) = \frac{s_2(s_l - s_1)}{s_l(s_2 - s_1)}\left(\frac{j - k_{i_2 j}}{s_2} + \rho_{i_2} - \left(\frac{j - k_{i_1 j}}{s_1} + \rho_{i_1}\right)\right)$$

and thus

$$k_{i_l j} = \frac{k_{i_1 j}(s_l - s_2) + k_{i_2 j}(s_l - s_1)}{s_2 - s_1} + s_l(\rho_{i_l} - \rho_{i_1}) + \frac{s_2(s_l - s_1)(\rho_{i_1} - \rho_{i_2})}{s_2 - s_1}.$$

This shows which elements of B' correspond to a monochromatic column. Recall the set B given in (2). Hence, the number of monochromatic columns is

$$N(q,D) = \sum_{(\beta_1,\dots,\beta_r)\in B} \prod_{l=1}^{r} \prod_{i=i_{l-1}+1}^{i_l} \frac{p_i - s_l}{q} + \left\lfloor \frac{s_l - (\beta_l + s_l(\rho_i - \rho_{i_l}) - 1) \bmod q - 1}{q} \right\rfloor + 1. \qquad \square$$

Proposition 3. *Assume $n > 1$ and $q \mid p_n$. Let $D' = D \setminus \{(p_n, \rho_n)\}$. Then*

$$N(q, D) = \frac{p_n}{q} N(q, D').$$

Proof. Let $P = p_1 p_2 \cdots p_n$. Note,

$$a_{ij} = \left(\left\lfloor \frac{j-1}{p_i} \right\rfloor + \rho_i \right) \bmod q$$

since $\lfloor (j-1)/p_i \rfloor$ calculates the number of complete blocks that the element is away from the beginning of the row. This number of blocks is added to the starting color of the row, ρ_i, and then taking this modulo q gives the color.

Let $1 \le x, y \le n-1$, $1 \le j \le P$, and $0 \le \alpha \le q-1$. Since $q \mid (P/p_x)$ and $q \mid (P/p_y)$, again the computations of (1) hold and we have

$$a_{x,j} - a_{x,\alpha P+j} \equiv a_{y,j} - a_{y,\alpha P+j} \pmod q.$$

This shows that if the first $n-1$ entries of column j are the same color then the first $n-1$ entries of column $j + \alpha P$ are the same color. Next, it is shown that $|\{a_{nj}, a_{n,P+j}, \dots, a_{n,(q-1)P+j}\}| = q$. Note that $(P/p_n) \not\equiv 0 \pmod q$. Now,

$$\begin{aligned} a_{nj} - a_{n,\alpha P+j} &\equiv \left\lfloor \frac{j-1}{p_n} \right\rfloor - \left\lfloor \frac{\alpha P + j - 1}{p_n} \right\rfloor \pmod q \\ &\equiv \left\lfloor \frac{j-1}{p_n} \right\rfloor - \left\lfloor \frac{j-1}{p_n} \right\rfloor + \frac{\alpha P}{p_n} \pmod q \\ &\equiv \frac{\alpha P}{p_n} \pmod q. \end{aligned}$$

Since q is prime, every color is represented in the set

$$\{a_{nj}, a_{n,P+j}, \dots, a_{n,(q-1)P+j}\}.$$

Therefore, $N(q, D) = (p_n/q) N(q, D')$. $\square$

3. Monochromatic column in three colors

In [Srivastava and Szabo 2008], there is a small issue in the results when 2 is one of the coprimes. The issue is that there is a possibility that β may only need to run up to 2 instead of 3. This can be seen in the general results of the previous section. We make the corrections while also restating the results with our simplified notation. Throughout this section, let n be a positive integer, $D = \{(p_i, \rho_i)\}_{i=1}^n$, where $p_1, \dots, p_n$ are pairwise coprime positive integers and $\rho_i \in \{0, 1, 2\}$, and $A = (a_{ij})$ be the $(3, D)$-color matrix. The first result is a direct application of Proposition 1 for $q = 3$.

Proposition 4 [Srivastava and Szabo 2008, Lemma 1]. *Let $s \in \{1, 2\}$. Assume $p_i \equiv s \pmod 3$ for $i \in \{1, \dots, n\}$. Then*

$$N(3, D) = q \sum_{\beta=1}^{\min\{p_1,3\}} \prod_{i=1}^{n} \frac{p_i - s}{3} + \left\lfloor \frac{s - (\beta + s(\rho_i - \rho_1) - 1) \bmod 3 - 1}{3} \right\rfloor + 1.$$

Proposition 5 [Srivastava and Szabo 2008, Lemma 2]. *Assume*

$$\{p_i \bmod 3 \mid i \in I\} = \{1, 2\}.$$

Let $i_0 = 0$, $i_2 = n$, and $i_1$be such that, for $i \in I$, we have $p_i \equiv l$ for $i_{l-1} < i \le i_l$. Then

$$N(3, D) = \sum_{\beta_1=1}^{3} \sum_{\beta_2=1}^{\min\{p_n,3\}} \prod_{l=1}^{2} \prod_{i=i_{l-1}+1}^{i_l} \frac{p_i - l}{q} + \left\lfloor \frac{l - (\beta_l + l(\rho_i - \rho_{i_l}) - 1) \bmod q - 1}{q} \right\rfloor + 1.$$

Proof. In Proposition 2, if $r = 2$ then $B = B'$. Furthermore, when $p_i \equiv 1 \pmod 3$, we have $p_i > 3$. This result then follows. □

For completeness, the following result is included as well.

Proposition 6 [Srivastava and Szabo 2008, Lemma 3]. *Assume $n > 1$ and $3 \mid p_n$. Let $D' = D \setminus \{(p_n, \rho_n)\}$ and $A' = (3, D')$. Then*

$$N(3, D) = \frac{p_n}{3} N(3, D').$$

References

[Crowell 2016] L. Crowell, *The monochromatic column problem: the prime case*, master's thesis, Easter Kentucky University, 2016, available at https://encompass.eku.edu/etd/356/.

[Jiang et al. 1999] T. Jiang, P. Kearney, and M. Li, "Open problems in computational molecular biology", *ACM SIGACT News* **30**:3 (1999), 43–49.

[Nagpaul and Jain 2002] S. R. Nagpaul and S. K. Jain, "Columns of uniform color in a rectangular array with rows having cyclically repeated color patterns", *Discrete Math.* **254**:1-3 (2002), 371–392. MR Zbl

[Srivastava and Szabo 2008] A. K. Srivastava and S. Szabo, "The monochromatic column problem", *Discrete Math.* **308**:17 (2008), 3906–3916. MR Zbl

Received: 2019-07-05 Revised: 2019-08-06 Accepted: 2019-08-12

loran.crowell@gmail.com *Eastern Kentucky University, Richmond, KY, United States*

steve.szabo@eku.edu *Eastern Kentucky University, Richmond, KY, United States*

involve

msp.org/involve

INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2019 is US $195/year for the electronic version, and $260/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

involve

2019 vol. 12 no. 8

1944-4176(2019)12:8;1-Y