An interval approach to recognition of numerical matrices

Alexander Prolubnikov

Omsk State University

e-mail: `a.v.prolubnikov@mail.ru`

The statement of the problem

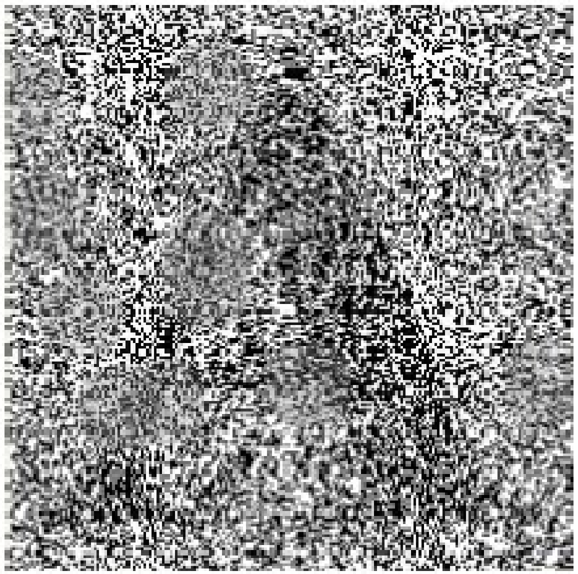
- $\{\mathcal{A}^{(k)}\}_{k=1}^{N}$ — pattern matrices, $a_{ij}^{(k)} \in \mathbb{R}$
- $\mathcal{A}$ is obtained in the process of noising from $\mathcal{A}^{(p)}$

  p is unknown
- it's known that matrices elements
  may be changed in the intervals

$$[a_{ij}^{(k)} - \Delta, a_{ij}^{(k)} + \Delta], \ \Delta > 0$$

We must define p (noised pattern matrix $\mathcal{A}^{(p)}$)

A E



Constructing of heuristics

We associate the input matrices with interval matrices:

$$\{\mathcal{A}^{(k)}\}_{k=1}^{N} \to \{\boldsymbol{A}^{(k)}\}_{k=1}^{N}$$

$\boldsymbol{A}^{(k)} = (\boldsymbol{a}_{ij}^{(k)})$:

$$\boldsymbol{a}_{ij}^{(k)} = [\min\{a_{ij}^{(k)}, a_{ij}\}, \max\{a_{ij}^{(k)}, a_{ij}\}],$$

where $a_{ij}^{(k)}$ are elements of pattern matrix $\mathcal{A}^{(k)}$,
a_{ij} are elements of recognized matrix $\mathcal{A}$

Elements of $\boldsymbol{A}^{(k)}$ are intervals,

which characterise changes of elements of pattern matrix $\mathcal{A}^{(k)}$ needed to obtain recognized matrix $\mathcal{A}$

Constructing of heuristics

Construct the systems of interval linear equations:

$$\boldsymbol{A}^{(k)}x = b, \;\; b \in \mathbb{R}^n$$

Suggestion: the lesser the variation of solutions of systems of linear equations, which gives the interval system $\boldsymbol{A}^{(k)}x = b$, the likely the recognized matrix $\mathcal{A}$ is obtained from $\mathcal{A}^{(k)}$

The variation of solutions of systems of linear equations, which gives $\boldsymbol{A}^{(k)}x = b$, is measured by

Lebesgue measure of united soluton set $\Xi(\boldsymbol{A}^{(k)}, b)$:

$\mu(\Xi(\boldsymbol{A}^{(k)}, b))$ is depends on:

- mutual disposition of elements of the matrices.
- It depends contionuously on their changes.

The selection of the right-hand side vector of the system

$\boldsymbol{A}^{(k)} x = b$

- the right-hand side vector is a real vector

It gives more precize enclosure of united solution set

because such selection decreases the distance between $\Xi(\boldsymbol{A}^{(k)}, b)$ and its interval hull $\square\Xi(\boldsymbol{A}^{(k)}, b)$

- if $b = e = (1, \ldots, 1)^{\top}$, then all of the elements of input matrix accounting at equal measure at the process of recognition

---

Thus, we consider the following systems of interval linear equations:

$$\boldsymbol{A}^{(k)} x = e$$

Computational complexity of the recognition

$\Xi^{(k)} \stackrel{def}{=} \Xi(\boldsymbol{A}^{(k)}, e)$

The problem of calculating of $\mu(\Xi^{(k)})$ has an exponential complexity

$\boldsymbol{X}^{(k)}$ is an approximaion of $\square\Xi^{(k)}$

$\boldsymbol{X}^{(k)}$ is a box: $\boldsymbol{X}^{(k)} = ([\underline{x}_1^k, \overline{x}_1^k], \ldots, [\underline{x}_n^k, \overline{x}_n^k])^\top$,

such that $\square\Xi^{(k)} \subseteq \boldsymbol{X}^{(k)}$

$$\mu(\boldsymbol{X}^{(k)}) = (\overline{x}_1^k - \underline{x}_1^k) \cdot \ldots \cdot (\overline{x}_n^k - \underline{x}_n^k)$$

If $Encl$ is some algorithm for enclosing of united solution set, then

$$C(N, n, Encl) = O(N \cdot C_{Encl}(n))$$

If $C_{Encl}(n) = O(n^2)$, then we have an algorithm

with lowest order of complexity

for algorithms of solution of the considered problem.

Modifications of the input matrices

Interval of change:

$$[a_{ij}^{(k)} - \Delta, a_{ij}^{(k)} + \Delta], \; \Delta > 0$$

Modification:

$$a_{ij} := a_{ij} + \upsilon$$

$$a_{ij}^{(k)} := a_{ij}^{(k)} + \upsilon$$

$$(\upsilon > 0)$$

As a result:

decreasing of the ratio:

$$\frac{\Delta}{|a_{ij}^{(k)}|} \to \frac{\Delta}{|a_{ij}^{(k)} + \upsilon|}$$

---

if the ratio $\Delta/|a_{ij}^{(k)}|$ is small enough then recognition is possible

Modifications of the input matrices

1) $\boldsymbol{A}^{(k)} := \boldsymbol{A}^{(k)} + \upsilon \boldsymbol{E}, \quad \boldsymbol{E}_{ij} = [1, 1],\ i, j = \overline{1, n}$

2) $\boldsymbol{A}^{(k)} := \boldsymbol{A}^{(k)} + \boldsymbol{D}, \quad \boldsymbol{D}$ is diagonal interval matrix

$$\boldsymbol{D}_{ii} = [D, D]$$

$$D^{(k)} = 2 \max_{1 \le i \le n} \sum_{j \ne i} |(\boldsymbol{A}^{(k)})_{ij}|, \quad D = \max_{1 \le k \le N} D^{(k)}$$

As a result:

$\boldsymbol{A}^{(k)}$ are H-matrices

We may use interval Gauss-Seidel method for enclosing $\square \Xi^{(k)}$

The initial approximation:

$$\text{box}\ ([-B, B], \ldots, [-B, B])^\top,\ B = 1/[\upsilon(n-1)]$$

The algorithm

Input: $\{\mathcal{A}^{(k)}\}_{k=1}^{N}$ and $\mathcal{A}$.
Output: Index p (matrix $\mathcal{A}^{(p)} \in \{\mathcal{A}^{(k)}\}_{k=1}^{N}$)

1. Construct matrices $\{\boldsymbol{A}^{(k)}\}_{k=1}^{N}$.

2. Using $Encl$ calculate $\boldsymbol{X}^{(k)}$, $k = \overline{1, N}$.

($\boldsymbol{X}^{(k)}$ are enclosures of $\Xi^{(k)}$)

3. Chose p such that $\mu(\boldsymbol{X}^{(p)}) = \min\limits_{1 \leq k \leq N} \mu(\boldsymbol{X}^{(k)})$.

p is a result of recognition

Total computational complexity:

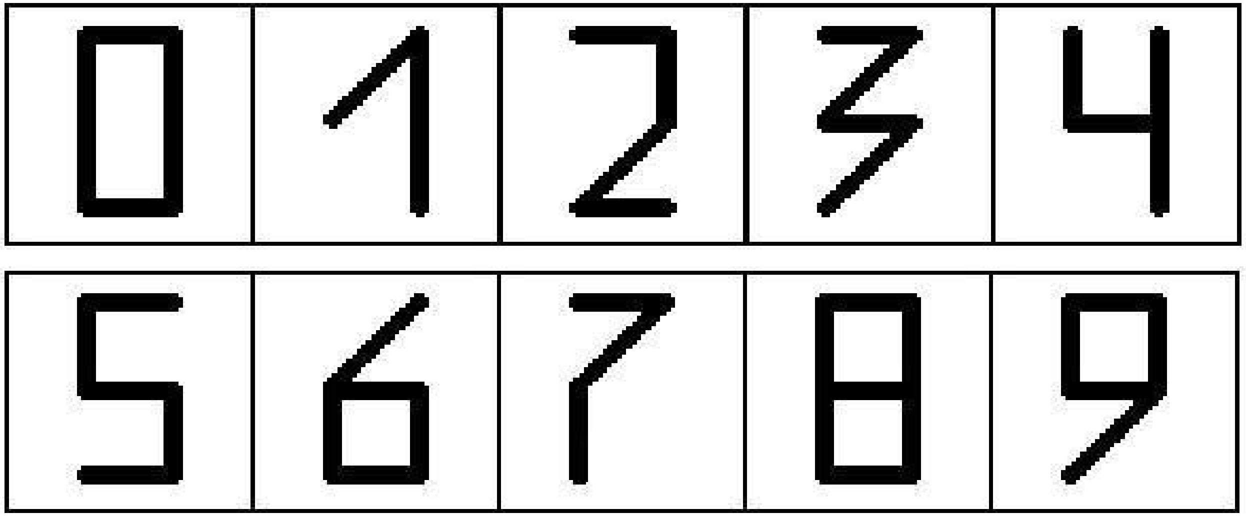
$Encl = GS$,

$$\underline{C(N, n, GS) = O(N \cdot N_{GS} \cdot n^2)}$$

$(N_{GS} = 20)$

Computational experiment

Pattern images



20×20, 35×35, 50×50 and 100×100 pixels resolution

$$a_{ij}^{(k)} = \begin{cases} c_1, & \text{if pixel in } ij \text{ position is white,} \\ c_2, & \text{if pixel in } ij \text{ position is black} \end{cases}$$

- black and white images: $c_1 = 0$ and $c_2 = 1$,
- greyscale images: $c_1,\ c_2 \in [0,\ 255]$

Level of noise $Q \in [0, 100]\,(\%)$

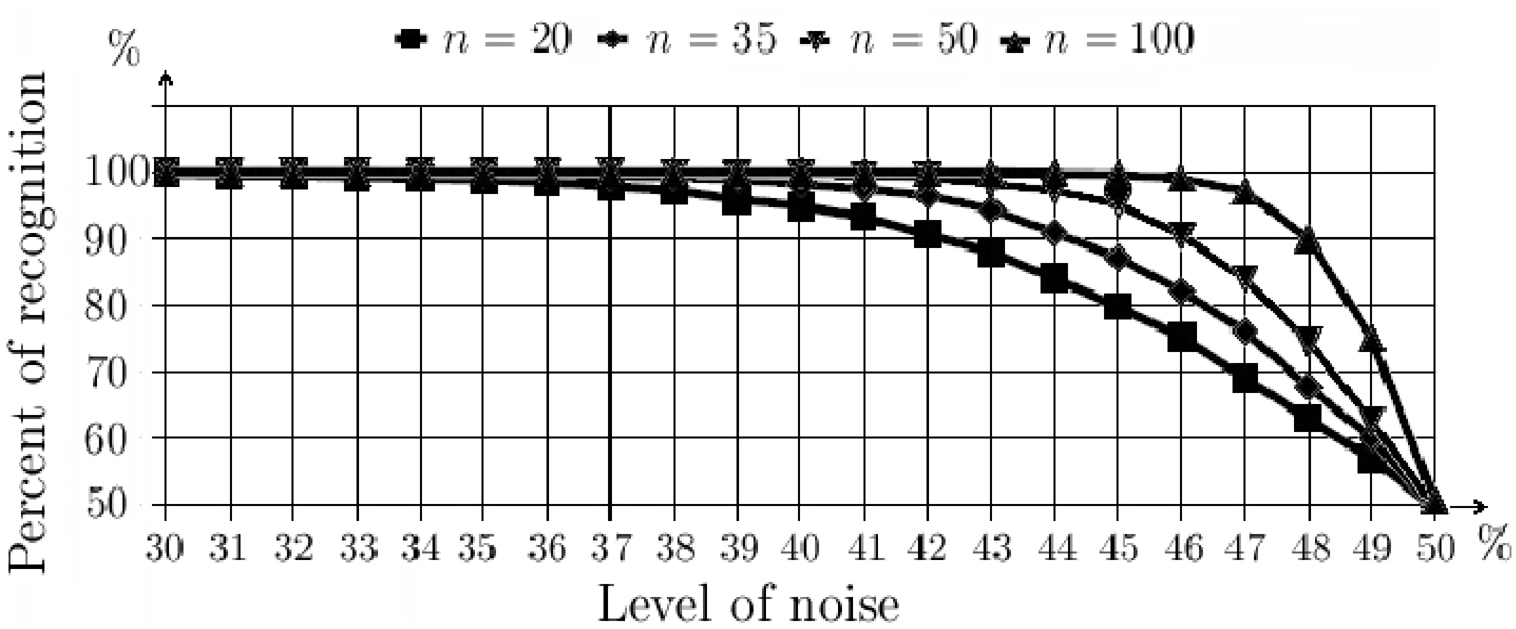
Percent of recognition

$$P = \frac{\text{number of correct recognition}}{\text{number of trials}} \times 100$$

Computational experiment

Q	31	32	33	34	35	36	37	38	39	40
$n = 20$	99.57	99.6	99.37	99.19	99	98.56	98	97.27	95.97	94.96
$n = 35$	99.9	99.97	99.97	99.84	99.83	99.71	99.47	99.36	99.02	98.31
$n = 50$	100	100	99.99	99.98	99.97	99.99	99.93	99.89	99.82	99.74
$n = 100$	100	100	100	100	100	100	100	100	100	100
Q	41	42	43	44	45	46	47	48	49	50
$n = 20$	93.32	90.67	88.07	84.1	79.87	75.28	69.16	63.06	57.02	49.72
$n = 35$	97.52	96.55	94.46	91	87.22	82.18	76.16	67.86	60.13	49.41
$n = 50$	99.51	99.27	98.36	97.29	95.23	90.6	83.99	74.61	62.69	49.86
$n = 100$	100	99.98	99.97	99.86	99.68	99.08	97.23	99.08	75.09	50.68

Percent of recognition for level of noise from 31% up to 50%, $c_1 = 0,\ c_2 = 1$



Computational experiment

Comparison of recognition efficiency of the presented heuristics with recognition efficiency of minimization of the distance $\rho(\mathcal{A}, \mathcal{A}^{(k)})$

$$\rho(\mathcal{A}, \mathcal{A}^{(k)}) = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(a_{ij} - a_{ij}^{(k)}\right)^2}$$

$$\rho(\mathcal{A}, \mathcal{A}^{(1)}) < \rho(\mathcal{A}, \mathcal{A}^{(2)}),$$

$$\text{but } |a_{ij} - a_{ij}^{(2)}| < |a_{ij} - a_{ij}^{(1)}|$$

for majority of ij positions of this matrices

Computational experiment

Comparison of recognition efficiency of the presented heuristics with recognition efficiency of minimization of the distance $\rho(\mathcal{A}, \mathcal{A}^{(k)})$

S is percent of the trials in which presented approach gives a recognition and miminimizing of $\rho(\mathcal{A}, \mathcal{A}^{(k)})$ doesn't give a recognition

$S, \%$	0	5.4	7.4	16.2	23.5
$P, \%$	100	99.93	99.79	99.72	99.81
Δ	10	25	50	75	100

Values of S when level of noise is equal to 44%, $c_1 = 110$, $c_2 = 120$

$S, \%$	22.8	37.5	47.3	46.4	46.4
$P, \%$	99.71	99.6	99.8	99.72	99.82
Δ	10	25	50	75	100

Values of S when level of noise is equal to 44%, $c_1 = 119$, $c_2 = 120$

Conclusions

- an algorithm of recognition of numerical matrices presented

- minimization of Lebesgue measure of united solution sets is the heuristics which the algorithm uses

- the recognition algorithm doesn't have a learning stage and it has a quadratic computational complexity