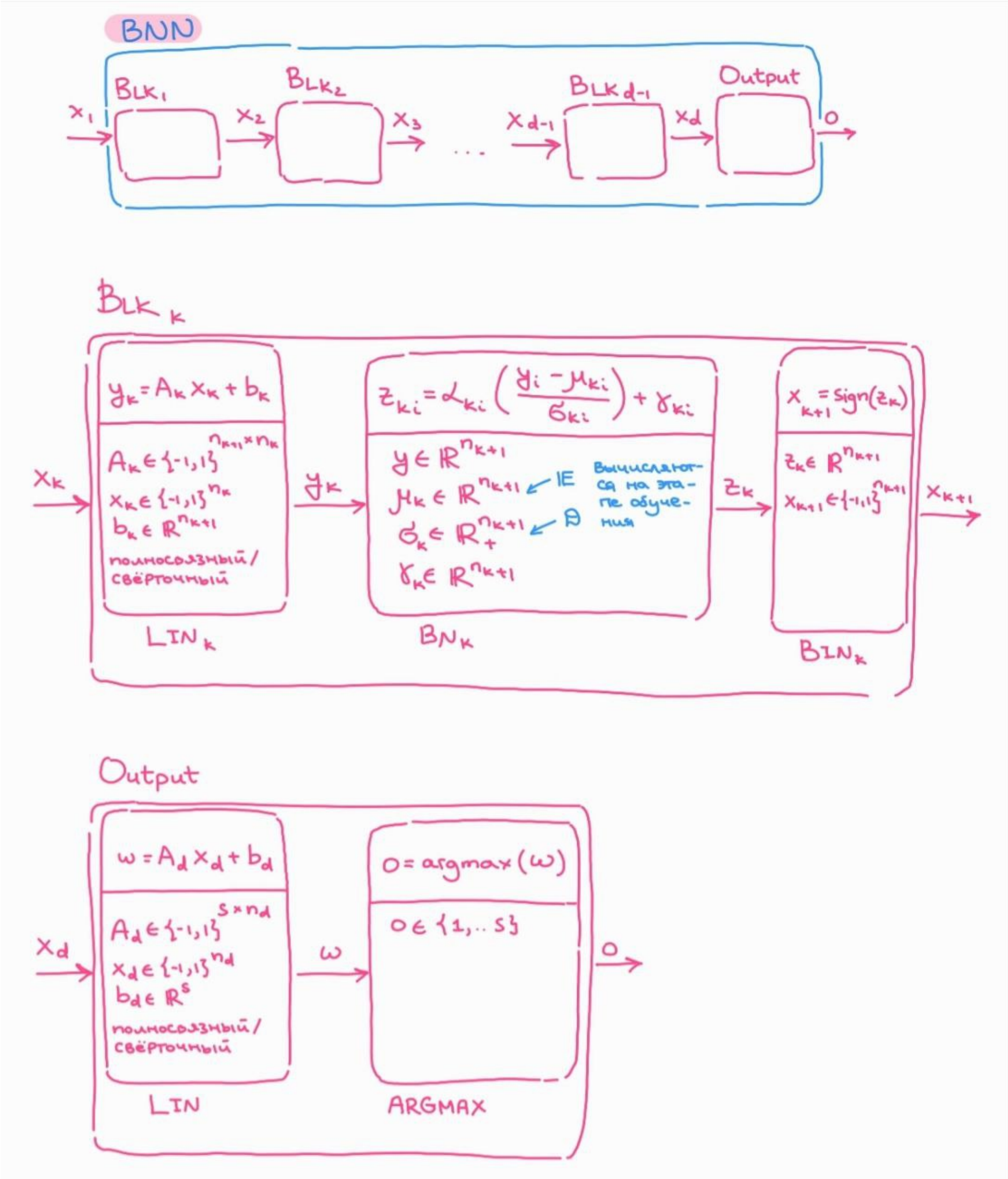
Структура BNN



Кодирование BNN в статье Народицкой

Алгоритм кодирования

1) Кодирование B_{LK_k}, в итоге получаем: $B_{IN}B_{LK_k}(x_k, x_{k+1})$

   а) MILP

   - L_{IN_k}

     Пусть a_i – i-ая строка A_k
     $y_i = <a_i, x_k> + b_i,\ i = 1 \ ...\ n_{k+1}$
     $y = (y_1, ...\ y_{n_{k+1}}) \in R^{n_{k+1}}$

   - B_{N_k}

     $z_{k_i} = \alpha_{k_i}(\frac{y_i - \mu_{k_i}}{\sigma_{k_i}}) + \gamma_{k_i}$

   - B_{IN_k}

     $\upsilon_i = sign(z_{k_i})$
     $sign(0) = 1$
     $z_{k_i} \geq 0 \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = -1$
     $x_{k+1} = (\upsilon_1, ...\ , \upsilon_{n_{k+1}})$

   б) ILP

   $\alpha_{k_i}(\frac{(<a_i, x_k> + b_i) - \mu_{k_i}}{\sigma_{k_i}}) + \gamma_{k_i} \geq 0 \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = -1$

   $\frac{\alpha_{k_i}}{\sigma_{k_i}}(<a_i, x_k> + b_i) - \frac{\alpha_{k_i}}{\sigma_{k_i}}\mu_{k_i} + \gamma_{k_i} \geq 0 \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = -1$
   $<a_i, x_k> \geq -\frac{\sigma_{k_i}}{\alpha_{k_i}}\gamma_{k_i} + \mu_{k_i} - b_i$

   Если $\alpha_{k_i} > 0$, то $<a_i, x_k> \geq \lceil -\frac{\sigma_{k_i}}{\alpha_{k_i}}\gamma_{k_i} + \mu_{k_i} - b_i \rceil$
   Если $\alpha_{k_i} < 0$, то $<a_i, x_k> \geq \lfloor -\frac{\sigma_{k_i}}{\alpha_{k_i}}\gamma_{k_i} + \mu_{k_i} - b_i \rfloor$

   Обозначим округление $(-\frac{\sigma_{k_i}}{\alpha_{k_i}}\gamma_{k_i} + \mu_{k_i} - b_i)$ как c_i

   Итого $<a_i, x_k> \geq c_i \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = -1$
   $x_{k+1} = (\upsilon_1, ...\ , \upsilon_{n_{k+1}})$

   в) SAT

   $\sum_{j=1}^{n_k} a_{i,j}x_{k,j} \geq c_i \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = -1$
   $x_{k+1} = (\upsilon_1, ...\ , \upsilon_{n_{k+1}})$

2) Кодирование $Output$, в итоге получаем: $B_{IN}O(x_d, o)$

   а) MILP

   - L_{IN}

     Пусть a_i – i-ая строка A_d
     $w_i = <a_i, x_d> + b_i,\ i = 1\ ...\ s$

$w = (w_1, \ldots w_s) \in R^s$

- $ARGMAX$

  $w_i \geq w_j \Leftrightarrow d_{ij} = 1,\ i, j = 1, \ldots, s$
  $\sum_{j=1}^{s} d_{ij} = s \Rightarrow o = i,\ i = 1, \ldots, s$

б) ILP

$<a_i, x_d> + b_i \geq <a_j, x_d> + b_j \Leftrightarrow d_{ij} = 1,\ i, j = 1, \ldots, s$
$<a_i - a_j, x_d> \geq b_j - b_i \Leftrightarrow d_{ij} = 1,\ i, j = 1, \ldots, s$
$<a_i - a_j, x_d> \geq \lceil b_j - b_i \rceil \Leftrightarrow d_{ij} = 1,\ i, j = 1, \ldots, s$
$\sum_{j=1}^{s} d_{ij} = s \Rightarrow o = i,\ i = 1, \ldots, s$

При этом известно, что $\sum_{j=1}^{s} d_{ij} \leq s \Rightarrow o = i,\ i = 1, \ldots, s$, значит можем переписать
$\sum_{j=1}^{s} d_{ij} \geq s \Rightarrow o = i,\ i = 1, \ldots, s$

в) SAT

$\sum_{l=1}^{n_d} (a_{i,l} - a_{j,l}) x_{d,l} \geq \lceil b_j - b_i \rceil \Leftrightarrow d_{ij} = 1,\ i, j = 1, \ldots, s$
$\sum_{j=1}^{s} d_{ij} \geq s \Rightarrow o = i,\ i = 1, \ldots, s$

3) Кодирование BNN: $(\wedge_{k=1}^{d-1} B_{IN} B_{LK_k}(x_k, x_{k+1})) \wedge B_{IN} O(x_d, o)$

Пример

Пусть в BNN два слоя: B_{LK} и $Output$.

- B_{LK}

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \text{ – входной вектор, } \widetilde{x} = \begin{pmatrix} \widetilde{x}_1 \\ \widetilde{x}_2 \\ \widetilde{x}_3 \end{pmatrix} \text{ – выходной вектор,}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 & 1 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 7.4 \\ 0.2 \\ -0.7 \end{pmatrix},$$

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.1 \\ -0.7 \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} 0.6 \\ -0.1 \\ -0.2 \end{pmatrix}, \sigma = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} -0.9 \\ 0.2 \\ -0.4 \end{pmatrix}, \gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.3 \\ -0.6 \end{pmatrix}$$

Кодирование:

$\alpha_1 = 0.1 > 0$
$\alpha_2 = 0.1 > 0$

$\alpha_3 = -0.7 < 0$

$$c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \lceil -\frac{\sigma_1}{\alpha_1}\gamma_1 + \mu_1 - b_1 \rceil \\ \\ \lceil -\frac{\sigma_2}{\alpha_2}\gamma_2 + \mu_2 - b_2 \rceil \\ \\ \lfloor -\frac{\sigma_3}{\alpha_3}\gamma_3 + \mu_3 - b_2 \rfloor \end{pmatrix} = \begin{pmatrix} \lceil -\frac{-0.9\cdot 0.1}{0.1} + 0.6 - 7.4 \rceil \\ \\ \lceil -\frac{0.2\cdot 0.3}{0.1} - 0.1 - 0.2 \rceil \\ \\ \lfloor -\frac{-0.4\cdot(-0.6)}{-0.7} - 0.2 + 0.7 \rfloor \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \\ 0 \end{pmatrix}$$

$$\widetilde{x} = \begin{pmatrix} \widetilde{x}_1 \\ \widetilde{x}_2 \\ \widetilde{x}_3 \end{pmatrix} = \begin{pmatrix} \text{Если } {<}a_1, x{>} \geq c_1\text{, то 1, иначе -1} \\ \\ \text{Если } {<}a_2, x{>} \geq c_2\text{, то 1, иначе -1} \\ \\ \text{Если } {<}a_3, x{>} \geq c_3\text{, то 1, иначе -1} \end{pmatrix} =$$

$$= \begin{pmatrix} \text{Если } a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + a_{15}x_5 \geq c_1\text{, то 1, иначе -1} \\ \\ \text{Если } a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 + a_{25}x_5 \geq c_2\text{, то 1, иначе -1} \\ \\ \text{Если } a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 + a_{35}x_5 \geq c_3\text{, то 1, иначе -1} \end{pmatrix} =$$

$$= \begin{pmatrix} \text{Если } x_1 + x_2 - x_3 - x_4 + x_5 \geq -5\text{, то 1, иначе -1} \\ \\ \text{Если } x_1 - x_2 + x_3 + x_4 + x_5 \geq 0\text{, то 1, иначе -1} \\ \\ \text{Если } -x_1 + x_2 - x_3 - x_4 + x_5 \geq 0\text{, то 1, иначе -1} \end{pmatrix}$$

- *Output*

$\widetilde{x} = \begin{pmatrix} \widetilde{x}_1 \\ \widetilde{x}_2 \\ \widetilde{x}_3 \end{pmatrix}$ – входной вектор, o – выход $\in \{1, 2\}$

$$\widetilde{A} = \begin{pmatrix} \widetilde{a}_{11} & \widetilde{a}_{12} & \widetilde{a}_{13} \\ \widetilde{a}_{21} & \widetilde{a}_{22} & \widetilde{a}_{23} \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}, \widetilde{b} = \begin{pmatrix} \widetilde{b}_1 \\ \widetilde{b}_2 \end{pmatrix} = \begin{pmatrix} -0.5 \\ 0.8 \end{pmatrix}$$

Кодирование:

$d_{ij} = 1 \Leftrightarrow \sum_{l=1}^{3}(a_{il} - a_{jl})x_{kl} \geq \lceil b_j - b_i \rceil \Leftrightarrow (a_{i1} - a_{j1})x_1 + (a_{i2} - a_{j2})x_2 + (a_{i3} - a_{j3})x_3 \geq \lceil b_j - b_i \rceil$

$D = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} = \begin{pmatrix} 1 & (2x_2 - 2x_3 \geq 2) \\ (-2x_2 + 2x_3 \geq -1) & 1 \end{pmatrix}$

$\sum_{j=1}^{s} d_{ij} \geq s \Rightarrow o = i,\ i = 1, \dots, s$

Кодирование BNN почти как в статье Ринарда

Все переменные $\in \{-1, 1\}$ теперь $\in \{0, 1\}$, таким образом можно кодировать свёрточные сети и есть разряженность

1) Кодирование B_{LK_k}, в итоге получаем: $B_{IN}B_{LK_k}(x_k, x_{k+1})$

а) MILP

- L_{IN_k}

  Пусть a_i – i-ая строка A_k
  $y_i = <a_i, x_k> + b_i, \ i = 1 \ ... \ n_{k+1}$
  $y = (y_1, ... \ y_{n_{k+1}}) \in R^{n_{k+1}}$

- B_{N_k}

  $z_{k_i} = \alpha_{k_i}(\frac{y_i - \mu_{k_i}}{\sigma_{k_i}}) + \gamma_{k_i}$

- B_{IN_k}

  $\upsilon_i = sign(z_{k_i})$
  $sign(0) = 1$
  $z_{k_i} \geq 0 \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = 0$
  $x_{k+1} = (\upsilon_1, ... \ , \upsilon_{n_{k+1}})$

б) ILP

$\alpha_{k_i}(\frac{(<a_i, x_k> + b_i) - \mu_{k_i}}{\sigma_{k_i}}) + \gamma_{k_i} \geq 0 \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = 0$

$\frac{\alpha_{k_i}}{\sigma_{k_i}}(<a_i, x_k> + b_i) - \frac{\alpha_{k_i}}{\sigma_{k_i}}\mu_{k_i} + \gamma_{k_i} \geq 0 \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = 0$
$<a_i, x_k> \geq -\frac{\sigma_{k_i}}{\alpha_{k_i}}\gamma_{k_i} + \mu_{k_i} - b_i$

Если $\alpha_{k_i} > 0$, то $<a_i, x_k> \geq \lceil -\frac{\sigma_{k_i}}{\alpha_{k_i}}\gamma_{k_i} + \mu_{k_i} - b_i \rceil$
Если $\alpha_{k_i} < 0$, то $<a_i, x_k> \geq \lfloor -\frac{\sigma_{k_i}}{\alpha_{k_i}}\gamma_{k_i} + \mu_{k_i} - b_i \rfloor$

Обозначим округление $(-\frac{\sigma_{k_i}}{\alpha_{k_i}}\gamma_{k_i} + \mu_{k_i} - b_i)$ как c_i

Итого $<a_i, x_k> \geq c_i \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = 0$
$x_{k+1} = (\upsilon_1, ... \ , \upsilon_{n_{k+1}})$

в) SAT

$\sum_{j=1}^{k} a_{i,j}x_{k,j} \geq c_i \Rightarrow \upsilon_i = 1$, иначе $\upsilon_i = 0$
$x_{k+1} = (\upsilon_1, ... \ , \upsilon_{n_{k+1}})$
Здесь в сумме стоят только 0 или 1. Каждое слагаемое – переменная, кодируем их в ограничения

2) Кодирование $Output$, в итоге получаем: $B_{IN}O(x_d, o)$

а) MILP

- L_{IN}

  Пусть a_i – i-ая строка A_d
  $w_i = <a_i, x_d> + b_i, \ i = 1 \ ... \ s$
  $w = (w_1, ... \ w_s) \in R^s$

- *ARGMAX*

  $w_i \geq w_j \Leftrightarrow d_{ij} = 1,\ i, j = 1, \dots, s$
  $\sum_{j=1}^{s} d_{ij} = s \Rightarrow o = i,\ i = 1, \dots, s$

б) ILP

$<a_i, x_d> + b_i \geq <a_j, x_d> + b_j \Leftrightarrow d_{ij} = 1,\ i, j = 1, \dots, s$
$<a_i - a_j, x_d> \geq b_j - b_i \Leftrightarrow d_{ij} = 1,\ i, j = 1, \dots, s$
$<a_i - a_j, x_d> \geq \lceil b_j - b_i \rceil \Leftrightarrow d_{ij} = 1,\ i, j = 1, \dots, s$
$\sum_{j=1}^{s} d_{ij} = s \Rightarrow o = i,\ i = 1, \dots, s$

При этом известно, что $\sum_{j=1}^{s} d_{ij} \leq s \Rightarrow o = i,\ i = 1, \dots, s$, значит можем переписать $\sum_{j=1}^{s} d_{ij} \geq s \Rightarrow o = i,\ i = 1, \dots, s$

в) SAT

$\sum_{l=1}^{n_d} (a_{i,l} - a_{j,l}) x_{d,l} \geq \lceil b_j - b_i \rceil \Leftrightarrow d_{ij} = 1,\ i, j = 1, \dots, s$

$\sum_{j=1}^{s} d_{ij} \geq s \Rightarrow o = i,\ i = 1, \dots, s$

Здесь тоже в сумме стоят только 0, 1 или -1, кодируем их как ограничения

3) Кодирование BNN: все ранее полученные ограничения

Пример

Пусть в BNN два слоя: B_{LK} и $Output$.

- B_{LK}

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \text{ – входной вектор, } \widetilde{x} = \begin{pmatrix} \widetilde{x}_1 \\ \widetilde{x}_2 \\ \widetilde{x}_3 \end{pmatrix} \text{ – выходной вектор,}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 7.4 \\ 0.2 \\ -0.7 \end{pmatrix},$$

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.1 \\ -0.7 \end{pmatrix}, \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} 0.6 \\ -0.1 \\ -0.2 \end{pmatrix}, \sigma = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} -0.9 \\ 0.2 \\ -0.4 \end{pmatrix}, \gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.3 \\ -0.6 \end{pmatrix}$$

Кодирование:

$\alpha_1 = 0.1 > 0$
$\alpha_2 = 0.1 > 0$
$\alpha_3 = -0.7 < 0$

$$c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \lceil -\frac{\sigma_1}{\alpha_1}\gamma_1 + \mu_1 - b_1 \rceil \\ \\ \lceil -\frac{\sigma_2}{\alpha_2}\gamma_2 + \mu_2 - b_2 \rceil \\ \\ \lfloor -\frac{\sigma_3}{\alpha_3}\gamma_3 + \mu_3 - b_2 \rfloor \end{pmatrix} = \begin{pmatrix} \lceil -\frac{-0.9 \cdot 0.1}{0.1} + 0.6 - 7.4 \rceil \\ \\ \lceil -\frac{0.2 \cdot 0.3}{0.1} - 0.1 - 0.2 \rceil \\ \\ \lfloor -\frac{-0.4 \cdot (-0.6)}{-0.7} - 0.2 + 0.7 \rfloor \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \\ 0 \end{pmatrix}$$

$$\widetilde{x} = \begin{pmatrix} \widetilde{x}_1 \\ \widetilde{x}_2 \\ \widetilde{x}_3 \end{pmatrix} = \begin{pmatrix} (<a_1, x> \geq c_1) \\ \\ (<a_2, x> \geq c_2) \\ \\ (<a_3, x> \geq c_3) \end{pmatrix} = \begin{pmatrix} (a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + a_{15}x_5 \geq c_1) \\ \\ (a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 + a_{25}x_5 \geq c_2) \\ \\ (a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 + a_{35}x_5 \geq c_3) \end{pmatrix} =$$

$$= \begin{pmatrix} (x_1 + x_2 + x_5 \geq -5) \\ \\ (x_1 + x_3 + x_4 + x_5 \geq 0) \\ \\ (x_2 + x_5 \geq 0) \end{pmatrix}$$

Из этого получаем ограничения:

Пусть 1, 2, 3, 4, 5 – номера входных переменных, а 6, 7, 8 – выходных, тогда

$1\ 2\ 5 \geq -5\ \#\ 6$
$1\ 3\ 4\ 5 \geq 0\ \#\ 7$
$2\ 5 \geq 0\ \#\ 8$

- *Output*

$\widetilde{x} = \begin{pmatrix} \widetilde{x}_1 \\ \widetilde{x}_2 \\ \widetilde{x}_3 \end{pmatrix}$ – входной вектор, o – выход $\in \{1, 2\}$

$\widetilde{A} = \begin{pmatrix} \widetilde{a}_{11} & \widetilde{a}_{12} & \widetilde{a}_{13} \\ \widetilde{a}_{21} & \widetilde{a}_{22} & \widetilde{a}_{23} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, $\widetilde{b} = \begin{pmatrix} \widetilde{b}_1 \\ \widetilde{b}_2 \end{pmatrix} = \begin{pmatrix} -0.1 \\ 0.8 \end{pmatrix}$

Кодирование:

$d_{ij} = 1 \Leftrightarrow \sum_{l=1}^{3} (\widetilde{a}_{il} - \widetilde{a}_{jl})\widetilde{x}_{kl} \geq \lceil \widetilde{b}_j - \widetilde{b}_i \rceil$

$D = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} = \begin{pmatrix} 1 & (\widetilde{x}_2 - \widetilde{x}_3 \geq 1) \\ (-\widetilde{x}_2 + \widetilde{x}_3 \geq 0) & 1 \end{pmatrix}$

$\sum_{j=1}^{s} d_{ij} \geq s \Rightarrow o = i,\ i = 1, \ldots, s$

Из этого получаем ограничения:

Пусть 9, 10, 11, 12 – номера переменных матрицы D, а 13, 14 – переменных, кодирующих итоговый результат, тогда

$7\ -8 \geq 1$ # 10
$-7\ 8 \geq 0$ # 11
$9\ 10 \geq 2$ # -13
$9\ 10 \geq 2$ # -14
$11\ 12 \geq 2$ # 13
$11\ 12 \geq 2$ # -14

Итого:

$1\ 2\ 5 \geq -5$ # 6
$1\ 3\ 4\ 5 \geq 0$ # 7
$2\ 5 \geq 0$ # 8
$7\ -8 \geq 1$ # 10
$-7\ 8 \geq 0$ # 11
$9\ 10 \geq 2$ # -13
$9\ 10 \geq 2$ # -14
$11\ 12 \geq 2$ # 13
$11\ 12 \geq 2$ # -14

MiniSatCS (Как я это вижу)

Пусть ограничения имеют вид: $y = (\sum_{i=1}^{n} l_i \geq b)$.

На этапе распространения добавились проверки:

1) Если хотя бы b переменных из l_i равны 1, то $y = 1$
2) Если хотя бы $n - b + 1$ переменных из l_i равны 0, то $y = 0$
3) Если $n - b$ переменных из l_i равны 1 и $y = 1$, то остальные l_i равны 1
4) Если $b - 1$ переменная из l_i равна 1 и $y = 0$, то остальные l_i равны 0