Noname manuscript No.
(will be inserted by the editor)

$\mathbb{Q}$-Bounds Consistency for the SPREAD Constraint with Variable Mean

Soon Chee Loong · Wen-Yang Ku · J. Christopher Beck

Received: date / Accepted: date

Abstract The SPREAD constraint enforces a relationship amongst a set of variables, their mean, and their standard deviation. The $\mathbb{Q}$-bounds consistency (BC) algorithms that have been formally published and the implementations of which we are aware all assume a fixed mean value. A sketch of the BC algorithm with variable mean was proposed, which relies on the continuity property of a key function used in the fixed mean case. We show that this function may be piecewise discontinuous, meaning that the extension of the algorithm to the variable mean case that is suggested in the literature is unsound. We propose a simple modification of the algorithm that achieves $\mathbb{Q}$-BC with variable mean.

1 Introduction

The SPREAD constraint enforces a given mean μ and maximum standard deviation σ among a set of n variables $\{X_1, \dots, X_n\}$. It can be defined as follows:

$$\texttt{spread}(\{X_1, \dots, X_n\}, \mu, \sigma),$$

where $\sigma = \sqrt{\sum_{i=1}^{n}(X_i - \mu)^2/n}$ and $\mu = \sum_{i=1}^{n} X_i/n$.

Bounds consistency for the SPREAD constraint is defined separately for continuous and integer domains as follows [4]:

Soon Chee Loong
Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada
E-mail: cheeloong.soon@mail.utoronto.ca

Wen-Yang Ku
E-mail: wku@mie.utoronto.ca

J. Christopher Beck
E-mail: jcb@mie.utoronto.ca

Definition 1 Let $I_D^{\mathbb{Q}}(X_i) = [X_i^{lb}, X_i^{ub}]$ be the rational interval domains of X_i, and let $I_D^{\mathbb{Z}}(X_i) = \{X_i^{lb}, \ldots, X_i^{ub}\}$ be the integer interval domains of X_i. A SPREAD constraint is $\mathbb{Q}$-bounds consistent (resp. $\mathbb{Z}$-bounds consistent) with respect to the domains of $\{X_1, \ldots, X_n\}$ if for all $i \in \{1, \ldots, n\}$ and each value $v_i \in \{X_i^{lb}, X_i^{ub}\}$, there exists values $v_j \in I_D^{\mathbb{Q}}(X_j)$ (resp. $v_j \in I_D^{\mathbb{Z}}(X_j)$) for all $j \in \{1, \ldots, n\} \setminus \{i\}$ such that $\texttt{spread}(\{X_1 = v_1, \ldots, X_n = v_n\}, \mu, \sigma)$ holds.

In other words, in $\mathbb{Z}$-BC, the support vector must be integer while in $\mathbb{Q}$-BC the support vector may not be.

The SPREAD constraint was originally proposed by Pesant et al. [2], where a $\mathbb{Q}$-BC algorithm that filters the domains of the X_i variables w.r.t μ and σ was presented. A simplified $\mathbb{Q}$-BC algorithm was presented in Schaus et al. [3,4], where an extended algorithm that filters the mean variable was proposed. The $\mathbb{Z}$-BC algorithm was introduced in Schaus et al. [4], demonstrating stronger filtering than $\mathbb{Q}$-BC. Both of these existing BC algorithms address the case where μ is fixed. This is a natural restriction in the applications that have been studied, for example, where a fixed amount of work is to be divided as evenly as possible amongst a set of workers.

For achieving domain consistency (DC), Pesant [1] proposed a pseudo-polynomial time filtering algorithm for the fixed mean case, where experiments were performed on three applications with fixed mean. The generalization to the variable mean case was discussed in Section 2.2 in Pesant [1, p. 695], showing that DC with variable mean can be achieved at a computational cost compared to the fixed mean case.

However, problems where μ is a variable do not seem to have been extensively studied, even though we believe there are compelling applications. For example, in a scheduling application where the objective is related to flow time (i.e., the time between when a job enters the system and when it is finished), it may be important to minimize some combination of the mean flow-time and its standard deviation: we want to achieve high-throughput but also ensure some measure of fairness amongst the jobs. Pesant [1,2] also notes that the variable mean case may arise in staff rostering problems.

The only reference to the variable mean case for BC we have found is in Schaus et al. [3], where the proposed extension to the fixed-mean $\mathbb{Q}$-BC algorithm depends on the claim that a particular function is concave and derivable. We show that this claim is incorrect and may lead to unsound inference. For the rest of the paper, we deal only with $\mathbb{Q}$-BC and modify the existing bounds propagation algorithm to properly address the variable mean case.

2 Recalling SPREAD Propagation with Fixed μ

We first present the propagation algorithm due to Schaus et al. [3,4] for fixed mean.

Following Schaus et al. [3], we define the necessary notation as follows. Let $I(X)$ be the set of intervals defined by pairs of consecutive elements of the sorted sequence of bounds of the variables $\{X_1, \dots, X_n\}$, we denote the k-th interval of $I(X)$ by I_k. Let $R(I_k) = \{X_i | X_i^{lb} \geq \max(I_k)\}$ be the set of variables whose values must be greater than I_k. In a number line representation, $R(I_k)$ are the variables whose values will necessarily lie to the right of I_k. Similarly, $L(I_k) = \{X_i | X_i^{ub} \leq \min(I_k)\}$ is the set of variables that lie to the left of I_k. We let $M(I_k)$ be the remaining variables not in $R(I_k)$ or $L(I_k)$ and let $m = |M(I_k)|$. We further let $ES(I_k)$ be the sum of the assigned X_i values at their extremes in $R(I_k)$ and $L(I_k)$, i.e., $ES(I_k) = \sum_{X_i \in R(I_k)} X_i^{lb} + \sum_{X_i \in L(I_k)} X_i^{ub}$.

Let $q = n\mu = \sum_{i=1}^{n} X_i$. It follows that $n\sigma^2 = \sum_{i=1}^{n} X_i^2 - \frac{q^2}{n}$. We can compute $V(I_k) = [ES(I_k) + (m)\min(I_k), ES(I_k) + (m)\max(I_k)]$, which is the interval of possible q values related to I_k. Let $I^q \in I(X)$ such that $q \in V(I^q)$. That is, I^q is the interval that contains q. To find the minimum consistent value of σ, all the variables in $R(I^q)$ should be assigned to their minimum value, X_i^{lb}. Similarly, all the variables in $L(I^q)$ should be assigned to their maximum value, X_i^{ub}. These are, in both cases, the domain values closest to the mean, which, as noted, lies in the interval I^q. All the remaining unassigned variables in $M(I^q)$ must be assigned the value, v, to satisfy the mean, after all the X_i variables in both $R(I^q)$ and $L(I^q)$ are assigned to their appropriate extreme value. This means we need to satisfy $mv + ES(I^q) = q$. Therefore we have $v = \frac{q - ES(I^q)}{m}$, and

$$\begin{aligned} n\sigma^2 &= \sum_{X_i \in R(I^q)} (X_i^{lb})^2 + \sum_{X_i \in L(I^q)} (X_i^{ub})^2 + mv^2 - \frac{q^2}{n} \qquad (1) \\ &= \sum_{X_i \in R(I^q)} (X_i^{lb})^2 + \sum_{X_i \in L(I^q)} (X_i^{ub})^2 + \frac{(q - ES(I^q))^2}{m} - \frac{q^2}{n}. \end{aligned}$$

When the mean is fixed, $\mu_{lb} = \mu_{ub}$, and q is a constant. We can locate the interval I^q such that $V(I^q)$ contains q and classify all the X_i variables to the sets $R(I^q)$, $L(I^q)$, and $M(I^q)$. We restrict our analysis to $R(I^q)$, as the derivation for $M(I^q)$ can be reduced to the same procedure as for the case $R(I^q)$, and the derivation for $L(I^q)$ is symmetric.

Pruning $X_i^{ub}\ \forall\ X_i \in R(I^q)$. The pruning of X_i has to respect the current upper bound on the standard deviation variable. Let σ_{shift} be the standard deviation that results from adding d to X_i^{lb}, where $d > 0$. Let d_{max} be the shift amount in X_i^{lb} such that $\sigma_{shift} = \sigma_{ub}$. Any shift larger than d_{max} will result in $\sigma_{shift} > \sigma_{ub}$, which renders the assignment inconsistent. Therefore, we can prune X_i using $X_i = [X_i^{lb}, min(X_i^{lb} + d_{max}, X_i^{ub})]$. Shifting X_i^{lb} by d shifts v to $\frac{q - (ES(I^q) + d)}{m} = v - \frac{d}{m}$. From Equation 1 we have

$$n\sigma_{shift}^2 = \sum_{X_j \in L(I^q)} (X_i^{ub})^2 + \sum_{X_j \in R(I^q), X_j \neq X_i} (X_i^{lb})^2 + (X_i^{lb} + d)^2 + m(v - \frac{d}{m})^2 - \frac{q^2}{n}.$$

Therefore, $n\sigma^2_{shift}$ is a function of d that can be written as follows:

$$n\sigma^2_{shift}(d) = d^2(1 + \frac{1}{m}) + 2d(X_i^{lb} - v) + n\sigma^2. \quad (2)$$

To solve for d_{max}, let Equation 2 equal to $n\sigma^2_{ub}$ and solve for d, we have

$$n\sigma^2_{shift}(d) = d^2(1 + \frac{1}{m}) + 2d(X_i^{lb} - v) + n\sigma^2 = n\sigma^2_{ub}.$$

$$d^2(1 + \frac{1}{m}) + 2d(X_i^{lb} - v) + n\sigma^2 - n\sigma^2_{ub} = 0.$$

Therefore, using the quadratic formula,

$$d_{max} = -b' + \frac{\sqrt{b'^2 - ac}}{a}, \quad (3)$$

where $a = 1 + \frac{1}{m}$, $b = 2(X_i^{lb} - v)$, $c = n\sigma^2 - n\sigma^2_{ub}$, $b' = \frac{b}{2}$.

Note that if $d_{max} > q - \min V(I^q)$, the v value does not reside in I^q. In this case, d_{max} is not valid and it needs to be calculated recursively through the intervals. We refer to Schaus [3,4] for further details.

3 A Problem with Propagation with a Variable μ

In the variable mean case, for each X_i, we have to compute $q^{max} \in [q^{lb} = n\mu_{lb}, q^{ub} = n\mu_{ub}]$ such that d_{max} is maximized, as different $q \in [q^{lb}, q^{ub}]$ result in different d_{max}. We can only take the largest d_{max}, as any other value may result in invalid pruning and thus potentially remove solutions. After q^{max} is computed, the propagation on X is performed in the same way as the fixed-mean case with q^{max}. In the original paper [3, p. 72] d_{max} is expressed as a function of q and it is stated that since $d_{max}(q)$ "can be shown to be concave and derivable, one can search [for] a q^0 such that d_{max} is maximum: $\frac{\partial d_{max}(q)}{\partial q}|_{q=q^0} = 0$." If q^0 does not reside within $[q^{lb}, q^{ub}]$, we can take the maximum of q^{lb} and q^{ub}, as $d_{max}(q)$ is concave.

The above statement is incorrect as it assumes that $d_{max}(q)$ is continuous throughout the interval $[q^{lb}, q^{ub}]$. Since $[q^{lb}, q^{ub}]$ can be spread across different $V(I^i)$s and thus be divided by different values of $a = 1 + \frac{1}{m}$, $d_{max}(q)$ may be divided into discontinuous intervals. Since $d_{max}(q)$ may be different for each discontinuous interval, using a single $d_{max}(q)$ may result in errors as we may not find the largest d_{max}.

We first show that the $d_{max}(q)$ function is concave, derivable at each interval but may not be piecewise continuous. In such a case, the partial derivative with respect to q is not well-formed.

Lemma 1 *$d_{max}(q)$ is concave, derivable within each interval but may not be piecewise continuous.*

Proof In order to determine the concavity of $d_{max}(q)$, we first reason on the signs of coefficients of $d_{max}(q)$. From Equation (3), since $a = 1 + \frac{1}{m}$, we know that $a > 0$. Since $X_i^{lb} - v \geq 0$, it follows that $b = 2(X_i^{lb} - v) \geq 0$. $c \leq 0$ as $n\sigma^2_{shift}(d) \leq n\sigma_{ub}$.

We discuss the two cases for c as follows:

- $c = 0$: We have $d_{max}(q) = \frac{(-b'+|b'|)}{a} = 0$. This means d_{max} is linear and thus concave.
- $c < 0$: From Equation (2), we know that $n\sigma^2_{shift}(d) \geq 0$ and it is convex. Therefore $c = n\sigma^2_{shift}(d) - n\sigma_{ub}$ is also convex. Since $c < 0$, c is convex and it resides below 0. As $a > 0$, ac remains convex and $ac < 0$. Multiplying a function by -1 reflects it upon the x-axis, so $-ac$ is concave and $-ac > 0$. Adding a constant positive value only shifts the function upwards. Thus, since $b'^2 \geq 0$, $b'^2 - ac$ must also be concave. It follows that $(b'^2 - ac) \geq 0$. Taking the square root of a concave function only scales the graph at each point, but it remains continuous and concave. Thus, $\sqrt{(b'^2 - ac)} \geq 0$. Last, adding a constant does not change the concavity, thus, $-b' + \sqrt{(b'^2 - ac)}$ maintains its concavity and continuity.
  However, for each interval, the expression $-b' + \sqrt{(b'^2 - ac)}$ might be divided by a different value of a, as $a = 1 + \frac{1}{m}$ for $m \neq 0$. This means that for any two consecutive intervals with different m, $d_{max}(q)$ is discontinuous, i.e., the function is discontinuous at the boundary of the intervals j and $j+1$ when $m_j \neq m_{j+1}$, and continuous when $m_j = m_{j+1}$.

Figure 1 shows an example of the $d_{max}(q)$ function. Variables and domains are the same as Example 1 in Schaus et al. [3], and we set the upper bound $\sigma_{ub} = \sqrt{(8/3)}$ (different notation for the upper bound, π_1^{max}, was used in the original paper with the relationship $\pi_1^{max} = n\sigma^2_{ub}$). As each variable results in a different $d_{max}(q)$ function, Figure 1 shows the function of variable X_1. It is clear that only intervals 2 and 3 are continuous at their boundaries. In addition, the partial derivative of the $d_{max}(q)$ function with respect to q is never zero. Therefore, the algorithm needs to check the values on the boundaries of the intervals.

As a result, the original algorithm for computing d_{max} [3, p. 72] cannot be implemented, as $d_{max}(q)$ is not a continuous function.

4 Fixing the Propagation with a Variable μ

Fortunately, we can fix the propagation algorithm with a straightforward modification. Unlike the fixed mean case where all the X_i variables share the same q, we have to calculate different q_{max} values for each X_i. As each individual interval represents a continuous concave function, each interval is derivable within itself. Therefore, we have to find the maximum d_{max} and the corresponding q_{max} in each interval. After q_{max} is determined, we can use it to propagate X_i as for the $\mu_{lb} = \mu_{ub}$ case.

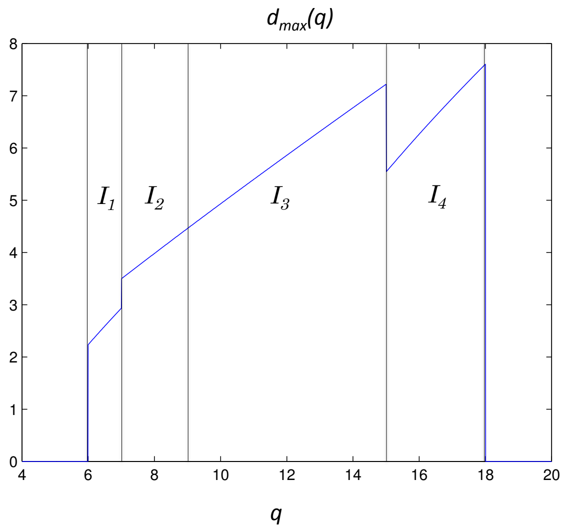

Fig. 1: The $d_{max}(q)$ function of variable X_1 of the example presented in Schaus et al. [3]. The x-axis represents q and the y-axis represents d_{max}. The $d_{max}(q)$ function is piecewise discontinuous. The q_{max} value can be found by executing Algorithm 1: Since the $d_{max}(q)$ function is increasing for all the intervals, we have to compute q^0 for each of the intervals. Since $q^0 \notin [q_j^{lb}, q_j^{ub}], \forall j$, we have $q_{max} = q_4^{ub} = 18$.

Without loss of generality, assume that q^{lb} and q^{ub} are already consistent with respect to all the X_i variables and σ. The details on filtering q, i.e., the filtering of the mean variable, are presented in Schaus et al. [3]. Let q_j^{lb} and q_j^{ub} be the smallest and largest q values, i.e., end points of interval I_j. The algorithm for computing d_{max} for one variable is presented in Algorithm 1. Note that X and $I(X)$ are the sets of all the variables and intervals.

To search for the maximum possible d_{max}, we iterate through each interval and update the q_{max} value. We use an additional variable dir to keep track of the direction of the $d_{max}(q)$ function at the current interval. If $d_{max}(q)$ is entirely decreasing in the current interval, we are sure that q^0 is outside of all the intervals afterwards. Therefore, we can set q_{max} to be either the current q_{max} or q_j^{lb}, depending which one gives a higher d_{max} value. If d_{max} is not entirely decreasing, we need to check if the current interval contains q^0 where $\frac{\partial d_{max}(q)}{\partial q}|_{q_0} = 0$. The two possibilities are as follows.

- Case 1: The first derivative q^0 resides in the current interval, i.e., $q^0 \in [q_j^{lb}, q_j^{ub}]$. In this case, this interval is the only interval that contains its first derivative. As $d_{max}(q)$ is concave, all the intervals to the left of this interval have maximum values at their upper bounds q_j^{ub}, and similarly, all the intervals to the right of this interval have maximum values at their lower

Algorithm 1 $FindDMaxVarMean(X, I(X))$

Data: X, $I(X)$
Results: d_{max} s.t. $\sigma_{shift} = \sigma_{ub}$ with $X_i^{ub} = X_i^{lb} + d_{max}$
Initialization: $q_{max} = q_1^{lb}$, dir = up
if $d_{max}(q_1^{lb}) > d_{max}(q_1^{ub})$ **then**
  dir = down
end if
for $1 \leq j \leq |I(X)|$ **do**
  if dir = down **then**
    $q_{max} = q_{max}$ if $d_{max}(q_{max}) > d_{max}(q_j^{lb})$, else $q_{max} = q_j^{lb}$
  else
    compute q^0
    if $q^0 \in [q_j^{lb}, q_j^{ub}]$ **then**
      $q_{max} = q_{max}$ if $d_{max}(q_{max}) > d_{max}(q^0)$, else $q_{max} = q^0$
      dir = down
    end if
    $q_{max} = q_{max}$ if $d_{max}(q_{max}) > d_{max}(q_j^{ub})$, else $q_{max} = q_j^{ub}$
  end if
end for
$d_{max} = FindDMax(X, I(q_{max}))$

bounds q_{lb}^j. In other words, since we already exclude the case that $d_{max}(q)$ is decreasing, q_{max} is equal to either the current q_{max} or q^0, depending on which one gives a higher d_{max} value.

– Case 2: The first derivative q^0 does not reside in the current interval, which means that $d_{max}(q)$ is still going upwards in the current interval. Therefore, we can set q_{max} to be either the current q_{max} or q_j^{ub}, depending which one gives a higher d_{max} value.

Since d_{max} represents the maximum shift that can occur to X_i^{lb}. By using q values other than q_{max}, the resulting d_{max} value may be less than the true value. This means that pruning X_i^{ub} may remove a feasible or optimal solution.

The complexity to enforce bounds consistency with fixed mean is $O(n\log(n))$ time, as d_{max} is computed in $O(\log(n))$ time for each of the n variables [4]. The complexity for the variable mean case thus increases to $O(n^2)$ time, since q_{max} is computed in $O(n)$ time.

5 Conclusion and Future Work

We have presented a propagation algorithm for the SPREAD constraint with a variable mean. Specifically, we have shown the piecewise discontinuous property of the specific function used to derive bounds reduction. Our proposed algorithm ensures $\mathbb{Q}$-BC, which cannot be guaranteed with the straightforward extension to the existing algorithm suggested in the literature.

One interesting extension is to develop an algorithm to achieve $\mathbb{Z}$-BC for the variable mean case. It is also valuable to implement the proposed algorithm and solve real world applications with variable mean.

References

1. Pesant, G.: Achieving domain consistency and counting solutions for dispersion constraints. INFORMS Journal on Computing **27**(4), 690–703 (2015)
2. Pesant, G., Régin, J.C.: Spread: A balancing constraint based on statistics. In: Principles and Practice of Constraint Programming-CP 2005, pp. 460–474. Springer (2005)
3. Schaus, P., Deville, Y., Dupont, P., Régin, J.C.: Simplification and extension of the spread constraint. In: Third international workshop on constraint propagation and implementation, pp. 77–91 (2006)
4. Schaus, P., Régin, J.C.: Bound-consistent spread constraint. EURO Journal on Computational Optimization pp. 1–24 (2013)