Approximating the diagonal of a Hessian: which sample set of points should be used

Gabriel Jarry–Bolduc

Received: 23 April 2021 / Accepted: 15 March 2022 / Published online: 6 April 2022
© Crown 2022

Abstract
An explicit formula based on matrix algebra to approximate the diagonal entries of a Hessian matrix with any number of sample points is introduced. When the derivative-free technique called generalized centered simplex gradient is used to approximate the gradient, then the formula can be computed for only one additional function evaluation. An error bound is introduced and provides information on the form of the sample set of points that should be used to approximate the diagonal of a Hessian matrix. If the sample set of points is built in a specific manner, it is shown that the technique is a $O(\Delta_S^2)$ accurate approximation of the diagonal entries of the Hessian matrix, where $\Delta_S$ is the radius of the sample set.

Keywords (Generalized) centered simplex gradient · Centered simplex Hessian diagonal · Derivative-free optimization

1 Introduction

One of the applications of derivative-free optimization (DFO) is to construct optimization algorithms that do not employ first-order information within the algorithm. Recently, substantial progress has been made regarding their applications and numerical implementations (see [1, 2, 6, 8, 14, 17]).

One of the main categories of DFO algorithms is model-based DFO methods. A simple method to approximate the objective function is to build a linear interpolation model from $n + 1$ well-poised sample points in $\mathbb{R}^n$. The gradient of this linear model is called the simplex gradient and provides an approximation of the true gradient [4, 16].
An error bound comparing the simplex gradient and the true gradient was introduced in the late 1990s and it is of order $O(\Delta S)$, where $\Delta S$ is the radius of the sample set of evaluated points [16]. This error bound shows that the optimizer can control the accuracy of the approximation technique by varying the radius $\Delta S$ of the sample set of points [2, Ch. 10 & 11]. A generalization of the simplex gradient called \textit{generalized simplex gradient} has the advantage of not being limited to the setting where exactly $n + 1$ interpolation points are used in $\mathbb{R}^n$. In [5], the authors consider the case where less than $n + 1$ sample points (underdetermined models) and more than $n + 1$ points (overdetermined models) are used to approximate the gradient. Most importantly, they establish an error bound for the overdetermined case and show that it retains order $O(\Delta S)$. This topic is also investigated in [21] and calculus rules for the generalized simplex gradient are introduced in [11, 21].

Many other methods of approximating gradients exist [3, 19, 20, 22, 24, 26]. One of these methods is the \textit{centered simplex gradient} [16]. Using the properties of the Moore-Penrose pseudoinverse (see Definition 2.1), the approach is generalized so that it does not require exactly $2n$ sample points in $\mathbb{R}^n$. The approach is called \textit{generalized centered simplex gradient} (GCSG) and it is created by retaining the $k$ original points in the sample set and adding their reflection through the point of interest (see Definition 3.2). An error bound which applies to the underdetermined, determined and overdetermined cases is introduced in [12]. The error bound shows that the GCSG is $O(\Delta_2^3)$ accurate.

In [9], Custódio and Vicente presented a linear system that can be solved to obtain a simplex gradient and an approximation of the $n$ diagonal terms of a simplex Hessian using $2n + 1$ sample points. More recently, approximating a full Hessian matrix with (generalized) simplex gradients has been investigated in [13]. An explicit formula that works for any number of sample points is developed.

In 2021, Coope and Tappenden showed how to compute the GCSG in $O(n)$ flops when using four different sample sets of points [7]. For these specific choices of sample sets, a formula to approximate the diagonal entries of the Hessian matrix is provided. We can observe that if the gradient is approximated via the GCSG, then only one additional function evaluation is sufficient to obtain an approximation of the diagonal entries of the Hessian matrix. Numerical examples are provided to compare the accuracy of the gradient and diagonal entries of the Hessian matrix depending on the sample set of points utilized. These numerical examples agree with the error bound defined for the gradient, but no error bound for the diagonal entries of the Hessian matrix is provided. For this reason, the poor accuracy obtained using certain sample sets of points is not fully explained. In this paper, an error bound for the diagonal entries of the Hessian matrix is introduced. Analyzing this error bound, we will obtain valuable information on the form of the sample set that should be utilized. Moreover, an explicit formula to compute the diagonal entries of a Hessian matrix for any nonempty sample set of points will be introduced. The technique is called \textit{centered simplex Hessian diagonal} (CSHD). The formula is developed by using the results presented in [7], and by using the properties of the Moore-Penrose pseudoinverse. For this reason, this paper can be viewed as an extension of [7].

The structure of this paper is the following. In Section 2, we introduce notation and basic definitions. In Section 3, the formula for the CSHD is introduced and
an error bound is proven. When the sample set of points has a specific form, it is shown that the CSHD is $O(\Delta^2)$ accurate. In Section 4, numerical examples are provided for different sample sets of points and an error analysis is provided. Section 5 summarizes the work accomplished and suggests some topics to explore in future research.

2 Preliminaries

Unless otherwise stated, we use the standard notation found in [23]. The domain of a function $f$ is denoted by $\text{dom } f$. The transpose of a matrix $A$ is denoted by $A^\top$. We work in the finite-dimensional space $\mathbb{R}^n$ with inner product $x^\top y = \sum_{i=1}^n x_i y_i$. The norm of a vector is denoted $\|x\|$ and is taken to be the $\ell_2$ norm. Given a matrix $A \in \mathbb{R}^{n \times k}$, the $\ell_2$ induced matrix norm is used. That is $\|A\|_2 = \|A\| = \max \{\|Ax\| : \|x\| = 1\}$.

We denote by $B(x^0, \Delta)$ the closed ball centered at $x^0$ with radius $\Delta$. Let $M \in \mathbb{R}^{n \times n}$. We will use the notation $\text{diag}[M]$ to denote the vector in $\mathbb{R}^n$ containing the diagonal entries of $M$. The notation $\text{Diag}[M] \in \mathbb{R}^{n \times n}$ represents the matrix containing the diagonal entries of a matrix $M$. The identity matrix in $\mathbb{R}^{n \times n}$ is denoted by $\text{Id}_n$ and the vector of all ones in $\mathbb{R}^n$ is denoted by $\mathbf{1}_n$.

Recall that a generalization of the matrix inverse is the Moore-Penrose pseudoinverse.

Definition 2.1 (Moore-Penrose pseudoinverse) [15] Let $A \in \mathbb{R}^{n \times k}$. The Moore-Penrose pseudoinverse of $A$, denoted by $A^\dagger$, is the unique matrix in $\mathbb{R}^{k \times n}$ that satisfies the following four equations:

1) $AA^\dagger A = A$, 2) $A^\dagger AA^\dagger = A^\dagger$, 3) $(AA^\dagger)^\top = AA^\dagger$, 4) $(A^\dagger A)^\top = A^\dagger A$.

The Moore-Penrose inverse $A^\dagger$ satisfies the following three properties.

1) If $A$ has full column rank $k$, then $A^\dagger$ is a left-inverse of $A$, that is, $A^\dagger A = \text{Id}_k$.
2) If $A$ has full row rank $n$, then $A^\dagger$ is a right-inverse of $A$, that is, $AA^\dagger = \text{Id}_n$.
3) If $A$ is a square matrix with full rank, then $A^\dagger = A^{-1}$, the inverse of $A$.

Last, recall the definition of the Hadamard product.

Definition 2.2 [15] Let $A = [a_{i,j}] \in \mathbb{R}^{n \times k}$ and $B = [b_{i,j}] \in \mathbb{R}^{n \times k}$. The Hadamard product of $A$ and $B$, denoted $A \odot B$ is the component wise product. That is $A \odot B = [a_{i,j}b_{i,j}] \in \mathbb{R}^{n \times k}$.
3 The CSHD and its error bound

We now present a process to obtain a formula to approximate the diagonal entries of the Hessian matrix. The process is assuming that the gradient is approximated via the GCSG [12]. Essentially, the formula is obtained by following the same steps than the ones in [7, Section 2 & Section 4.2], and by using the properties of the Moore-Penrose pseudoinverse to allow an arbitrary number of sample points.

Let \( f : \text{dom} \, f \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \) be \( C^2 \), \( x^0 \in \text{dom} \, f \) be the point of interest, and \( S = [s^1 \ s^2 \ \ldots \ s^k] \in \mathbb{R}_{n \times k} \) be a set of nonzero distinct directions in \( \mathbb{R}^n \) written in matrix form. The matrix \( S \) contains all the directions to add and subtract to the point of interest \( x^0 \) to form the sample points where the function \( f \) is evaluated. Suppose a model \( m(x) \) is constructed such that

\[
f(x^0 + s^i) = m(x^0 + s^i), \quad i \in \{1, 2, \ldots, k\},
\]

where \( m \) is a diagonal quadratic model. That is

\[
m(x) = f(x^0) + g^\top (x - x^0) + \frac{1}{2} (x - x^0)^\top D(x - x^0)
\]

where \( D \approx \text{Diag} \left[ \nabla^2 f(x^0) \right] \in \mathbb{R}^{n \times n} \) and \( g \approx \nabla f(x^0) \in \mathbb{R}^n \). Define

\[
W = [s^1 \odot s^1 \ \ldots \ s^k \odot s^k] \in \mathbb{R}^{n \times k}, \quad \text{and} \quad \delta_f(x^0; S) = \begin{bmatrix} f(x^0 + s^1) - f(x^0) \\ \vdots \\ f(x^0 + s^k) - f(x^0) \end{bmatrix} \in \mathbb{R}^k.
\]

Letting \( x = x^0 + s^i, i \in \{1, 2, \ldots, k\} \), in (1), we get the system

\[
\delta_f(x^0; S) = S^\top g + \frac{1}{2} W^\top d
\]

where \( d = \text{diag} \left[ D(x^0) \right] \in \mathbb{R}^n \). When the GCSG is employed, the sample points \( x^0 - s^i \) for all \( i \in \{1, \ldots, k\} \) are also created. Letting \( x = x^0 - s^i \) for all \( i \) in (1), we obtain the system

\[
\delta_f(x^0; -S) = -S^\top g + \frac{1}{2} W^\top d.
\]

System (2) and System (3) can be combined to form one block matrix equation:

\[
\begin{bmatrix}
\delta_f(x^0; S) \\
\delta_f(x^0; -S)
\end{bmatrix} =
\begin{bmatrix}
S^\top & \frac{1}{2} W^\top \\
-S^\top & \frac{1}{2} W^\top
\end{bmatrix}
\begin{bmatrix}
g \\
d
\end{bmatrix}.
\]

System (4) is simplified by performing the following block row operations on Row 1 and Row 2 respectively: (1) Multiply Row 1 by \(-1\) and subtract Row 2, (2) multiply Row 2 by \(-1\) and add Row 1. The block matrix system is now

\[
\begin{bmatrix}
-\delta_f(x^0; S) - \delta_f(x^0; -S) \\
\delta_f(x^0; S) - \delta_f(x^0; -S)
\end{bmatrix} =
\begin{bmatrix}
0 & -W^\top \\
2S^\top & 0
\end{bmatrix}
\begin{bmatrix}
g \\
d
\end{bmatrix}.
\]

From (5), we find

\[
\frac{1}{2} \left( \delta_f(x^0; S) - \delta_f(x^0; -S) \right) = S^\top g
\]
\[ \delta_f(x^0; S) + \delta_f(x^0; -S) = W^T d. \]

Now let
\[ \delta^c_f(x^0; S) = \frac{1}{2} \begin{bmatrix} f(x^0 + s^1) - f(x^0 - s^1) \\ \vdots \\ f(x^0 + s^k) - f(x^0 - s^k) \end{bmatrix} \in \mathbb{R}^k \]

and
\[ \varepsilon_f(x^0; S) = \begin{bmatrix} f(x^0 + s^1) + f(x^0 - s^1) - 2f(x^0) \\ \vdots \\ f(x^0 + s^k) + f(x^0 - s^k) - 2f(x^0) \end{bmatrix} \in \mathbb{R}^k. \]

We obtain
\[ \frac{1}{2} \left( \delta_f(x^0; S) - \delta_f(x^0; -S) \right) = \delta^c_f(x^0; S) = S^T g \quad (6) \]

and
\[ \left( \delta_f(x^0; S) + \delta_f(x^0; -S) \right) = \varepsilon_f(x^0; S) = W^T d. \quad (7) \]

Analyzing (6), we see that if \( S \) is full row rank, then \((S^T)^\dagger\) is a left-inverse of \( S^T \) and we can solve (6) for \( g \) by premultiplying both sides by \((S^T)^\dagger\). Note that when \( S \) is a non-square matrix with full row rank, then the Moore-Penrose pseudoinverse provides the unique least squares solution of (6) [15, P. 453]. Similarly, if \( W \) is full row rank then \((W^T)^\dagger\) is a left-inverse of \( W^T \) and we can solve (7) for \( d \) by premultiplying both sides of (7) by \((W^T)^\dagger\). We are now ready to introduce an explicit formula to approximate the gradient and a formula to approximate the diagonal of the Hessian matrix based on (6) and (7).

**Definition 3.1** (Generalized centered simplex gradient) [12, Definition 2.7] Let \( f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \), \( x^0 \in \text{dom } f \) be the point of interest, \( S = [s^1 \ s^2 \ \cdots \ s^k] \) be in \( \mathbb{R}^{n \times k} \) and \( W = [s^1 \odot s^1 \ \cdots \ s^k \odot s^k] \in \mathbb{R}^{n \times k} \). Assume that \( x^0 \pm s^i \in \text{dom } f \) for all \( i \). The generalized centered simplex gradient of \( f \) at \( x^0 \) over \( S \) is denoted by \( \nabla_c f(x^0; S) \) and defined by
\[ \nabla_c f(x^0; S) = (S^T)^\dagger \delta^c_f(x^0; S) \in \mathbb{R}^n. \]

To be consistent with the notations used so far in this paper, the notation chosen for the GCSG in the previous definition is slightly different than the one used in [12].

**Definition 3.2** (Centered simplex Hessian diagonal) Let \( f : \text{dom } f \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \), \( x^0 \in \text{dom } f \) be the point of interest, \( S = [s^1 \ s^2 \ \cdots \ s^k] \) be in \( \mathbb{R}^{n \times k} \) and \( W = [s^1 \odot s^1 \ \cdots \ s^k \odot s^k] \in \mathbb{R}^{n \times k} \). Assume that \( x^0 \pm s^i \in \text{dom } f \) for all \( i \). The centered simplex Hessian diagonal of \( f \) at \( x^0 \) over \( S \), denoted by \( \text{diag} \left[ \nabla^2_c f(x^0; S) \right] \) is a vector in \( \mathbb{R}^n \) given by
\[ \text{diag} \left[ \nabla^2_c f(x^0; S) \right] = (W^T)^\dagger \varepsilon_f(x^0; S). \]
Note that the CSHD can be obtained for only one additional function evaluation when the GCSG has been already computed (by evaluating \( f \) at \( x^0 \)). If the function value at the point of interest \( x^0 \) is known, then the CSHD can be obtained for free in terms of function evaluations.

The formula in the previous definition involves the Moore-Penrose pseudoinverse of an \( n \times k \) real matrix. A computationally inexpensive approach to compute the Moore-Penrose pseudoinverse is to use the singular value decomposition of the matrix \([25]\). For instance, one of the most advanced methods, the Golub-Reinsch SVD algorithm, requires \( 4n^2k + 8nk^2 + 9k^3 \) flops \([10]\). Therefore, the previous formulae for the GCSG and CSHD cannot be computed directly in \( O(n) \) flops. The reader is referred to \([7]\) for more details on how to rewrite the formulae so that they can be computed in \( O(n) \) flops when \( S \) has a specific form.

Next, we introduce an error bound for the CSHD. Analyzing the error bound, we will get a better understanding of the matrix \( S \) that should be used to obtain an accurate approximation of the diagonal entries of the Hessian matrix. In the following theorem, the parameter \( \Delta_S \) is the radius of the matrix \( S = [s^1 \cdots s^k] \in \mathbb{R}^{n \times k} \). That is
\[
\Delta_S = \max_{i \in \{1, 2, \ldots, k\}} \| s^i \|.
\]

**Theorem 3.3** (Error bound for the CSHD) *Let \( f : \text{dom} \, f \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \) be \( C^4 \) on an open domain containing \( B(x^0, \Delta_S) \) where \( x^0 \in \text{dom} \, f \) is the point of interest and \( \Delta_S > 0 \) is the radius of \( S = [s^1 \cdots s^k] \in \mathbb{R}^{n \times k} \). Let \( W = S \circ S \in \mathbb{R}^{n \times k} \). Denote by \( L_{\nabla^3 f} \geq 0 \) the Lipschitz constant of \( \nabla^3 f \) on \( B(x^0, \Delta_S) \). If \( W \) has full row rank, then*
\[
\| \text{diag} \left[ \nabla^2 f(x^0; S) \right] - \text{diag} \left[ \nabla^2 f(x^0) \right] \| \\
\leq \left\| (\tilde{W})^\dagger \right\| \left( \frac{k}{12} L_{\nabla^3 f} \Delta_S^2 + 2 \sum_{i=1}^{k} \| (\tilde{s}^i)^\dagger U [\nabla^2 f(x^0)] \tilde{s}^i \| \right) 
\]

*where \( \tilde{W} = W / \Delta_S^2 \) and \( \tilde{s}^i = s^i / \Delta_S \) for all \( i \in \{1, 2, \ldots, k\} \).*

**Proof** Since \( W \) has full row rank, \( W^\dagger \) has full column rank and so \( (W^\dagger)^\dagger \) is a left-inverse of \( W^\dagger \). We obtain
\[
\| \text{diag} \left[ \nabla^2 f(x^0; S) \right] - \text{diag} \left[ \nabla^2 f(x^0) \right] \| \\
= \left\| (W^\dagger)^\dagger \varepsilon_f(x^0; S) - \text{diag} \left[ \nabla^2 f(x^0) \right] \| \\
\leq \left\| (W^\dagger)^\dagger \right\| \| \varepsilon_f(x^0; S) - W^\dagger \text{diag} \left[ \nabla^2 f(x^0) \right] \| .
\]

Let us investigate the vector \( \varepsilon_f(x^0; S) - W^\dagger \text{diag} \left[ \nabla^2 f(x^0) \right] \in \mathbb{R}^k \). Row \( i \) of this vector can be written as \( f(x^0 + s^i) + f(x^0 - s^i) - 2f(x^0) - (w^i)^\dagger \text{diag} \left[ \nabla^2 f(x^0) \right] \). By Taylor’s Theorem, we know
\[
f(x^0 + s^i) = f(x^0) + \nabla f(x^0)^\dagger s^i + \frac{1}{2} (s^i)^\dagger \nabla^2 f(x^0) s^i \\
+ (s^i)^\dagger \nabla^3 f(x^0)[s^i] s^i + R_3(x^0 + s^i) 
\]

\( \Box \) Springer
where \( R_3(x^0 + s^i) \) is the remainder term defined as in [18, Theorem 4.7] and \( \nabla^3 f(x^0)[s^i] \in \mathbb{R}^{n \times n} \) is a directional Hessian. That is
\[
\nabla^3 f(x^0)[s^i] = \lim_{\tau \to 0} \frac{\nabla^2 f(x^0 + \tau s^i) - \nabla^2 f(x^0)}{\tau}.
\]

Similarly, by Taylor's Theorem we may write
\[
f(x^0 - s^i) = f(x^0) - \nabla f(x^0)^\top s^i + \frac{1}{2} (s^i)^\top \nabla^2 f(x^0) s^i - (s^i)^\top \nabla^3 f(x^0)[s^i] s^i + R_3(x^0 - s^i).
\]

Adding (9) and (10) together and rearranging, we obtain
\[
f(x^0 + s^i) + f(x^0 - s^i) - 2 f(x^0) = (s^i)^\top \nabla^2 f(x^0) s^i + R_3(x^0 + s^i) + R_3(x^0 - s^i).
\]

Now, note that \( \nabla^2 f(x^0) = \text{Diag}[\nabla^2 f(x^0)] + N[\nabla^2 f(x^0)]. \)

We get
\[
f(x^0 + s^i) + f(x^0 - s^i) - 2 f(x^0) = (s^i)^\top N[\nabla^2 f(x^0)] s^i + R_3(x^0 + s^i) + R_3(x^0 - s^i).
\]

The matrix \( N[\nabla^2 f(x^0)] \) can be written as \( N[\nabla^2 f(x^0)] = U[\nabla^2 f(x^0)]' + U[\nabla^2 f(x^0)]'' \). Hence, we have \( (s^i)^\top N[\nabla^2 f(x^0)] s^i = 2 (s^i)^\top U[\nabla^2 f(x^0)] s^i \) for all \( i \). Therefore, we obtain the equation
\[
f(x^0 + s^i) + f(x^0 - s^i) - 2 f(x^0) = (w^i)^\top \text{diag}[\nabla^2 f(x^0)]
\]
\[
= 2 (s^i)^\top U[\nabla^2 f(x^0)] s^i + R_3(x^0 + s^i) + R_3(x^0 - s^i).
\]

It follows that
\[
\| e_f(x^0; S) - W^\top \text{diag}[\nabla^2 f(x^0)] \|
\leq \sum_{i=1}^{k} |2 (s^i)^\top U[\nabla^2 f(x^0)] s^i + R_3(x^0 + s^i) + R_3(x^0 - s^i)|
\leq \sum_{i=1}^{k} |2 (s^i)^\top U[\nabla^2 f(x^0)] s^i| + \sum_{i=1}^{k} |R_3(x^0 + s^i) + R_3(x^0 - s^i)|
\leq 2 \sum_{i=1}^{k} |(s^i)^\top U[\nabla^2 f(x^0)] s^i| + \sum_{i=1}^{k} \frac{L_{\nabla^3 f}}{12} \Delta^4 S
\leq 2 \sum_{i=1}^{k} |(s^i)^\top U[\nabla^2 f(x^0)] s^i| + \frac{k}{12} L_{\nabla^3 f} \Delta^4 S.
Finally, the absolute error is
\[
\| \text{diag} \left[ \nabla^2_c f(x^0; S) \right] - \text{diag}[\nabla^2 f(x^0)] \| \\
\leq \left\| (W^\top)^\dagger \left( 2 \sum_{i=1}^k |(s^i)^\top U[\nabla^2 f(x^0)]s^i| + \frac{k}{12} L_{\nabla^3 f} \Delta_S^2 \right) \right\|. \\
= \left\| (\tilde{W}^\top)^\dagger \left( 2 \sum_{i=1}^k |(\tilde{s}^i)^\top U[\nabla^2 f(x^0)]\tilde{s}^i| + \frac{k}{12} L_{\nabla^3 f} \Delta_S^2 \right) \right\|. \\
\]

Note that the error bound given in (8) does not necessarily need to go to zero as the radius \( \Delta_S \) tends to zero. Hence, the true absolute error does not need to go to zero as \( \Delta_S \to 0 \). To ensure that the true absolute error goes to zero when \( \Delta_S \to 0 \), we need the term \( \sum_{i=1}^k |(\tilde{s}^i)^\top U[\nabla^2 f(x^0)]\tilde{s}^i| \) to vanish. Let us introduce a type of matrix that guarantees that the term \( \sum_{i=1}^k |(\tilde{s}^i)^\top U[\nabla^2 f(x^0)]\tilde{s}^i| = 0 \).

**Definition 3.4** (Lonely matrix) The matrix \( S \in \mathbb{R}^{n \times k} \) is a lonely matrix if there is exactly one nonzero entry in each column of \( S \).

**Lemma 3.5** Let \( S = [s^1 \ s^2 \ \cdots \ s^k] \) be a lonely matrix in \( \mathbb{R}^{n \times k} \). Let \( U \in \mathbb{R}^{n \times n} \) be a strictly upper triangular matrix. Then
\[
\sum_{i=1}^k |(s^i)^\top Us^i| = 0. \\
\]

**Lemma 3.6** Let \( S = [s^1 \ s^2 \ \cdots \ s^k] \) be a lonely matrix in \( \mathbb{R}^{n \times k} \) with full row rank. Then \( W = S \odot S \) has full row rank.

Obviously, lemma 3.5 applies to the set \( S = \text{Id}_n \) as it is a diagonal matrix with nonzero entries. Lemma 3.6 shows that it is easy to guarantee that \( W \) is full row rank by simply taking \( S \) to be a lonely matrix with full row rank. The following corollary provides an error bound when \( S \) is a lonely matrix with full row rank.

**Corollary 3.7** Let \( f : \text{dom } f \subseteq \mathbb{R}^n \to \mathbb{R} \) be \( C^4 \) on an open domain containing \( B(x^0, \Delta_S) \) where \( x^0 \in \text{dom } f \) is the point of interest and \( \Delta_S > 0 \) is the radius of the matrix \( S = [s^1 \ s^2 \ \cdots \ s^k] \in \mathbb{R}^{n \times k} \). Let \( W = S \odot S \in \mathbb{R}^{n \times k} \). Denote by \( L_{\nabla^3 f} \geq 0 \) the Lipschitz constant of \( \nabla^3 f \) on \( B(x^0, \Delta_S) \). If \( S \) is a lonely matrix with full row rank, then
\[
\| \text{diag} \left[ \nabla^2_c f(x^0; S) \right] - \text{diag}[\nabla^2 f(x^0)] \| \leq \left\| (\tilde{W}^\top)^\dagger \right\| \frac{k}{12} L_{\nabla^3 f} \Delta_S^2. \\
\]
The previous error bound shows that the CSHD is a \( O(\Delta_S^2) \) accurate approximation of the diagonal entries of the Hessian matrix. Under these assumptions, the optimizer can control the accuracy of the approximation technique. Indeed, as \( \Delta_S \) tends to zero, the error bound goes to zero. Hence, the true absolute error needs to go
to zero as $\Delta S \to 0$. Since the error bound involves the Lipschitz constant $L_{\nabla^3 f}$, the CSHD is perfectly accurate when $f$ is a polynomial of degree less than 4.

In the next section, numerical examples are provided. We will use different matrices $S$ and different points of interest to approximate the CSHD. The relative errors obtained for each $S$ and each point of interest will be compared.

### 4 Numerical examples

In this section, we begin by conducting a numerical experiment on the Rosenbrock function $f : \mathbb{R}^2 \to \mathbb{R} : (y_1, y_2) \mapsto (1 - y_1)^2 + 100(y_2 - y_1^2)^2$. The CSHD is computed to approximate the diagonal of the Hessian matrix at two points: $x^1 = [1.1 \ 1.2^2 + 10^{-5}]^\top$, and $x^2 = [0.9 \ 0.81]^\top$. Four different matrices of directions are employed:

1. The coordinate basis in $\mathbb{R}^{n \times n}$ (CB): $CB = \text{Id}_n$.
2. A regular basis in $\mathbb{R}^{n \times n}$ (RB): $RB = \sqrt{n+1} \left( \text{Id}_n - \frac{1}{n} \left( 1 - \sqrt{\frac{1}{n+1}} \right) \text{Id}_n \right)$.
3. A coordinate minimal positive basis in $\mathbb{R}^{n \times (n+1)}$ (CMPB): $CMPB = \left[ \text{Id}_n - \text{Id}_n \right]$.
4. A regular minimal positive basis in $\mathbb{R}^{n \times (n+1)}$ (RMPB): $RMPB = [RB - RB \text{Id}_n]$.

Note that the previous four matrices are full row rank, but only CB is a lonely matrix. First, the point of interest $x^1$ is considered and the positive parameter $h^1$ is set to $10^{-3}$. The CSHD is computed using four different sets $S$: $h^1 CB$, $h^1 RB$, $h^1 CMPB$, and $h^1 RMPB$. The experiment is repeated with a new point of interest, $x^2 = [0.9 \ 0.81]^\top$. The positive parameter $h^2$ is set to $10^{-6}$. To compare the accuracy obtained from each set $S$ and each point of interest in $\{x^1, x^2\}$, the relative error will be used instead of the absolute error. This choice is justified by the fact that the true value of the diagonal entries of the Hessian matrix are large in this example. For a given point of interest $x^0$ and matrix $S$, the relative error for a function $f$ at a point $x^0$ over $S$ is defined as

$$\text{RE}_f(x^0; S) = \frac{\| \text{diag}[\nabla^2 f(x^0; S)] - \text{diag}[\nabla^2 f(x^0)]\|}{\| \text{diag}[\nabla^2 f(x^0)]\|}.$$ 

Table 1 provides the results obtained. This numerical experiment is the same as the one conducted in [7, Section 5], but using relative errors instead of absolute errors to compare the matrices of directions.

Analyzing Table 1, we observe that the sets RB and CMPB provide poor approximations of the diagonal entries of the Hessian matrix.

| Table 1 | Relative error of the diagonal entries of the Hessian matrix |
|--------|-----------------------------------------------------------|
| Set $T$ | $\text{RE}_f(x^1; h^1 T)$ | $\text{RE}_f(x^2; h^2 T)$ |
| CB     | $2.02 \times 10^{-7}$ | $1.18 \times 10^{-9}$ |
| RB     | $3.14 \times 10^{-1}$ | $3.74 \times 10^{-1}$ |
| CMPB   | $4.19 \times 10^{-1}$ | $4.99 \times 10^{-1}$ |
| RMPB   | $1.78 \times 10^{-7}$ | $3.39 \times 10^{-9}$ |
As a second experiment, let us investigate if it is theoretically possible to obtain a smaller relative error than $4.19 \times 10^{-1}$ when considering the point of interest $x^1$ and the matrix of directions $S = h \cdot \text{CMPB}$ for some value of $h > 0$. Using Maple 2020, the relative error of $f$ at $x^1$ over $h \cdot \text{CMPB}$ is defined as a function of $h$. The command $\text{Minimize}$ from the $\text{Optimization}$ package is used to solve this minimization problem. We obtain that the minimum of the relative error $RE_f(x^1; h \cdot \text{CMPB})$ is approximately $2.96 \times 10^{-1}$. We see that this value is not equal to zero which is the case when the matrix of directions $h \cdot \text{CB}$ is used (or any lonely matrix with full row rank).

This agrees with the error bound defined (8); note that the radius of the set $S = h \cdot \text{CMPB}$ is $\Delta_S = h \sqrt{2}$. Looking at the term $\sum_{i=1}^{n+1} |(\hat{s}^i)^\top U [\nabla^2 f(x^1)] \hat{s}^i|$ in (8), we see that

$$\sum_{i=1}^{3} |(\hat{s}^i)^\top U [\nabla^2 f(x^1)] \hat{s}^i| = h^2 |\hat{I}_2^\top U [\nabla^2 f(x^1)] \hat{I}_2|$$

$$= \frac{1}{2} |\hat{I}_2^\top U [\nabla^2 f(x^1)] \hat{I}_2|$$

$$= \frac{1}{2} \left[ \begin{array}{cc} 1 & 1 \\ \hline 1 & 0 \\ 0 & 0 \end{array} \right] \left[ \begin{array}{ccc} 0 & -440 \\ 0 & 1 \end{array} \right] \left[ \begin{array}{c} 1 \\ 1 \end{array} \right] = 220.$$  

Observe that, independently of the value $h$, the sum $\sum_{i=1}^{3} |(\hat{s}^i)^\top U [\nabla^2 f(x^1)] \hat{s}^i|$ is always equal to 220 when the matrix CMPB is used on this specific function $f$ at the point of interest $x^1$. Therefore, the limit of the error bound in (8) as $h$ tends to zero (and so $\Delta_S$ tends to zero) is greater than zero. Hence, the error bound does not go to zero as $h$ tends to zero and so nothing guarantees that the true absolute error goes to zero as $h \to 0$.

Based on the results obtained in Table 1, the matrix of directions RMPB seems to provide better results than CMPB and RB. The following example shows that this is not always the case.

Let $g : \mathbb{R}^2 \to \mathbb{R} : (y_1, y_2) \mapsto y_1^2 + y_2^2 + 4y_1y_2$. Let $x^3 = [1 \ 10^8]^\top$ and $x^4 = [1 \ 1]^\top$ be the points of interest and let $h = 10^{-3}$ be the shrinking parameter. Table 2 provides the relative errors obtained.

Table 2 shows that the accuracy of the CSHD when using the matrix of directions RMPB may be greatly affected by our choice of point of interest.

The previous numerical results and the error bound developed in (8) suggest that CB, or any lonely matrix with full row rank, are appropriate when approximating the diagonal entries of the Hessian matrix via the technique $\text{diag} \left[ \nabla^2 f(x^0; S) \right]$.

| Set $T$   | $RE_g(x^3; hT)$ | $RE_g(x^4; hT)$ |
|-----------|-----------------|-----------------|
| CB        | 0               | 0               |
| RB        | $2.49 \times 10^3$ | $2.50 \times 10^3$ |
| CMPB      | $3.33 \times 10^3$ | $3.33 \times 10^3$ |
| RMPB      | $2.36 \times 10^4$ | $1.41 \times 10^{-3}$ |
5 Conclusion

Based on the error bound developed in Theorem 3.3 and our numerical results, the coordinate basis is the best choice of the four sets tested to approximate the diagonal entries of the Hessian matrix via the technique $\text{diag}\left[\nabla^2_c f(x^0; S)\right]$. More generally, a set of directions that forms a lonely matrix with full row rank is appropriate to approximate the diagonal entries of the Hessian matrix. When this is the case, the CSHD is $O(\Delta_S^2)$ accurate.

The matrix of directions RMPB seems to provide high accuracy for certain functions at an adequate point of interest. Based on the numerical results obtained in Section 4, a reasonable conjecture is that the matrix of directions RMPB performs well when the off-diagonal entries of the Hessian matrix are not too large compared to the diagonal entries and each entry of the point of interest is not too large. It could be interesting to investigate this topic in depth and try to rigorously characterize the necessary conditions for the matrix of directions RMPB, or other matrices of directions which are not lonely matrices, to provide high accuracy.

It is worth mentioning that approximating the diagonal of the Hessian matrix and not the entire Hessian may be misleading when the Hessian of $f$ is not diagonally dominate. For example, consider the function $f(y_1, y_2) = \alpha y_1 y_2$ where $\alpha$ is a nonzero scalar. Then

$$\nabla^2 f(y_1, y_2) = \begin{bmatrix} 0 & \alpha \\ \alpha & 0 \end{bmatrix}.$$  

In this case, the diagonal of the Hessian matrix does not provide any information about the function $f$ even though $f$ is not equal to the zero function.

Future research directions could include testing a diagonal quadratic model in a model-based trust region method [2, Chapter 11] and comparing its performance to a (full) quadratic model. It may be valuable to compare these two approaches in terms of function evaluations necessary to obtain a solution within a certain accuracy.

The procedure and notions used to develop the CSHD and the associated error bound are very similar to the ones used in [13] to develop an explicit formula to approximate a full Hessian matrix and the associated error bound. For this reason, it could be valuable to verify if the CSHD and the associated error bound can be easily derived from the analysis for the full Hessian matrix performed in [13]. A reasonable conjecture is that the CSHD may be obtained from the diagonal entries of a generalization of the nested-set Hessian where the second matrix of directions, called $T$ in [13], is allowed to be different for all generalized simplex gradients involved in the formula to compute the nested-set Hessian.

Acknowledgements Jarry-Bolduc would like to acknowledge UBC for the funding received through the University Graduate Fellowship award.

Funding Jarry-Bolduc’s research is partially funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada, Discover Grant #2018-03865.

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.
Declarations

Conflict of interest  The author declares no competing interests.

References

1. Audet, C.: A survey on direct search methods for blackbox optimization and their applications. In: Mathematics without Boundaries: Surveys in Interdisciplinary Research, pp. 31–56. Springer, New York (2014)
2. Audet, C., Hare, W.: Derivative-Free and Blackbox Optimization. Springer Series in Operations Research and Financial Engineering. Springer, Switzerland (2017)
3. Billups, S., Larson, J., Graf, P.: Derivative-free optimization of expensive functions with computational error using weighted regression. SIAM J. Optim. 23(1), 27–53 (2013)
4. Bortz, D., Kelley, C.: The simplex gradient and noisy optimization problems. In: Computational Methods for Optimal Design and Control, pp. 77–90. Springer (1998)
5. Conn, A., Scheinberg, K., Vicente, L.: Geometry of sample sets in derivative-free optimization: polynomial regression and underdetermined interpolation. IMA J. Numer. Anal. 28(4), 721–748 (2008)
6. Conn, A., Scheinberg, K., Vicente, L.: Introduction to Derivative-free Optimization, vol. 8. SIAM, Philadelphia, PA (2009)
7. Coope, I., Tappenden, R.: Gradient and diagonal Hessian approximations using quadratic interpolation models and aligned regular bases. Numerical Algorithms, pp. 1–25 (2021)
8. Custódio, A., Scheinberg, K., Vicente, L.: Methodologies and software for derivative-free optimization. In: Advances and Trends in Optimization with Engineering Applications, Chapter 37, pp. 495–506. SIAM, Philadelphia, PA (2017)
9. Custódio, A., Vicente, L.: Using sampling and simplex derivatives in pattern search methods. SIAM J. Optim. 18(2), 537–555 (2007)
10. Golub, G., Van Loan, C. Matrix Computations, 3rd edn. Johns Hopkins University Press, USA (1996)
11. Hare, W., Jarry-Bolduc, G.: Calculus identities for generalized simplex gradients: Rules and applications. SIAM J. Optim. 30(1), 853–884 (2020)
12. Hare, W., Jarry-Bolduc, G., Planiden, C.: Error bounds for overdetermined and underdetermined generalized centred simplex gradients. IMA Journal of Numerical Analysis (2020)
13. Hare, W., Jarry-Bolduc, G., Planiden, C.: Hessian approximations. arXiv preprint arXiv: 2011.02584 (2020)
14. Hare, W., Nutini, J., Tesfamariam, S.: A survey of non-gradient optimization methods in structural engineering. Adv. Eng. Softw. 59, 19–28 (2013)
15. Horn, R., Johnson, C.: Matrix Analysis. Cambridge University Press, Cambridge (1990)
16. Kelley, C.: Iterative Methods for Optimization, vol. 18. SIAM, Philadelphia, PA (1999)
17. Larson, J., Menickelly, M., Wild, S.: Derivative-free optimization methods. Acta Numerica 28, 287–404 (2019)
18. Lax, P., Terrell, M.: Multivariable Calculus with Applications. Springer, Switzerland (2017)
19. Ouevray, R., Bierlaire, M.: Boosters: A derivative-free algorithm based on radial basis functions. Int. J. Model. Simul. 29(1), 26–36 (2009)
20. Powell, M.: Least Frobenius norm updating of quadratic models That satisfy interpolation conditions. Math. Program. 100(1), 183–215 (2004)
21. Regis, R.: The calculus of simplex gradients. Optim. Lett. 9(5), 845–865 (2015)
22. Regis, R., Shoemaker, C.: Constrained global optimization of expensive black box functions using radial basis functions. J. Glob. Optim. 31, 153–171 (2005)
23. Rockafellar, R., Wets, R.: Variational Analysis. Fundamental Principles of Mathematical Sciences. Springer, Berlin (1998)
24. Schonlau, M., Welch, W., Jones, D.: Global versus local search in constrained optimization of computer models. Lecture Notes Monograph Series, pp. 11–25 (1998)
25. Wang, G., Wei, Y., Qiao, S.: Generalized Inverses: Theory and Computations, vol. 53. Springer, Berlin (2018)
26. Wild, S., Shoemaker, C.: Global convergence of radial basis function trust region derivative-free algorithms. SIAM J. Optim. 21(3), 761–781 (2011)

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.