

Automatic Differentiation Using Complex and Hypercomplex Variables

Applying dual numbers using the Cauchy-Riemann matrix form

University of Texas at San Antonio

July, 2023

Cauchy-Reimann (CR) form

- Dual numbers can be employed using all real numbers with no imaginary parts using Cauchy-Reimann matrices of all real values. This may be useful for numerical algorithms/software when a dual library is not available.

Any dual number $a + b\epsilon$ can be rewritten in the matrix form $\begin{bmatrix} a & 0 \\ b & a \end{bmatrix}^\dagger$.

Thus, to differentiate a function using dual numbers, one computes $f(x + \epsilon) = f\left(\begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}\right)$

Then

$$\frac{df}{dx} = f\left(\begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}\right)_{21}$$

Where the subscript 21 indicates the first column second row of the resulting matrix.

[†]An equivalent alternate form is $\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$.

Closed-form example: $f(x) = x^2$

- Compute the derivative of $f(x) = x^2$ using the CR form for dual numbers.

$$f(x) = x^2$$
$$f(x + ih) = \begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}^2 = \begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix} \begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix} = \begin{bmatrix} x^2 & 0 \\ 2x & x^2 \end{bmatrix}$$

$$\frac{df}{dx} = f\left(\begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}\right)_{21} = 2x$$

Closed-form example: $f(x) = x^3$

- Compute the derivative of $f(x) = x^3$ using the CR form for dual numbers.

$$f(x) = x^3$$
$$f(x + ih) = \begin{bmatrix} x & -h \\ h & x \end{bmatrix}^3 = \begin{bmatrix} x & -h \\ h & x \end{bmatrix} \begin{bmatrix} x^2 & 0 \\ 2x & x^2 \end{bmatrix} \begin{bmatrix} x^3 & 0 \\ 3x^2 & x^3 \end{bmatrix}$$

$$\frac{df}{dx} = f\left(\begin{bmatrix} x & -h \\ h & x \end{bmatrix}\right)_{21} = 3x^2$$

Closed-form example: $f(x) = e^x$

- Compute the derivative of $f(x) = e^x$ using the CR form for dual numbers.

$$f(x) = e^x$$
$$f(x + \epsilon) = \text{expM}\left(\begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}\right) = \begin{bmatrix} e^x & 0 \\ e^x & e^x \end{bmatrix}$$

$$\frac{df}{dx} = f\left(\begin{bmatrix} x & -h \\ h & x \end{bmatrix}\right)_{21} = e^x$$

Note, **functions of matrices** must be used for the exponential operator NOT the standard exponential function.

Closed-form example: $f(x) = \sin(x)$

- Compute the derivative of $f(x) = \sin(x)$ using the CR form for dual numbers.

$$f(x) = \sin(x)$$
$$f(x + \epsilon) = \text{sinM}\left(\begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}\right) = \begin{bmatrix} \sin(x) & 0 \\ \cos(x) & \sin(x) \end{bmatrix}$$

$$\frac{df}{dx} = f\left(\begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}\right)_{21} = \cos(x)$$

Note, **functions of matrices** must be used for the sine operator NOT the standard sine function.

Closed-form example: $f(x) = \cos(x)$

- Compute the derivative of $f(x) = \cos(x)$ using the CR form for dual numbers.

$$f(x) = \cos(x)$$
$$f(x + \epsilon) = \text{cosM}\left(\begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}\right) = \begin{bmatrix} \cos(x) \cosh(h) & \sin(x) \sinh(h) \\ -\sin(x) \sinh(h) & \cos(x) \cosh(h) \end{bmatrix}$$

$$\frac{df}{dx} = f\left(\begin{bmatrix} x & 0 \\ 1 & x \end{bmatrix}\right)_{12} = -\sin(x) \frac{\sin(h)}{h} = -\sin(x)$$

Note, **functions of matrices** must be used for the cosine operator NOT the standard cosine function.

Functions of Matrices

- Functions of matrices¹ are not the same as the standard functions. For example,

$$\text{Exp} \begin{bmatrix} a & 0 \\ c & d \end{bmatrix} = \begin{bmatrix} e^a & 0 \\ \frac{c(e^a - e^d)}{a - d} & e^d \end{bmatrix}$$

$$\text{Sin} \begin{bmatrix} a & 0 \\ c & d \end{bmatrix} = \begin{bmatrix} \sin(a) & 0 \\ \frac{c(\sin(a) - \sin(d))}{a - d} & \sin(d) \end{bmatrix}$$

- Hence to use the CR form, one must have a language with good support of functions of matrices.

¹Higham, Nicholas J. Functions of matrices: theory and computation. Society for Industrial and Applied Mathematics, 2008.

Functions of matrices

- As discussed in the CTSE section, the dual variable CR form has superior numerical performance compared to the complex variable CR form.

Complex form: $\begin{bmatrix} x & -h \\ h & x \end{bmatrix}$, h small

$$\frac{df}{dx} \approx \frac{\text{sqr}tM\left(\begin{bmatrix} x & -h \\ h & x \end{bmatrix}\right)_{21}}{h}$$

$$f(x) = \sqrt{x}$$

Dual form: $\begin{bmatrix} x & 0 \\ h & x \end{bmatrix}$, h arbitrary

$$\frac{df}{dx} = \frac{\text{sqr}tM\left(\begin{bmatrix} x & 0 \\ h & x \end{bmatrix}\right)_{21}}{h}$$

