OPTIMAL RICCI MATRIX COMPUTATIONS.

1. Notation

Throughout the paper we use $\times$ to denote pointwise multiplication. Consequently, we also use $f^{\times 2}$ to denote pointwise powers. We also assume that operators are represented by matrices and functions are column vectors.

We will use M_i to denote a row of a matrix, and $M_{:i}$ to denote its column.

Suppose r is a row vector, c is a column vector and M is a matrix. We can perform pointwise operations between objects not having the same size. For example $c + M$ should be understood as the column is first promoted to a square matrix with the same column repeated. The same way we can promote a row to a matrix, or a number to any shape. Note that such extension of $+$ and $-$ preserves distributive property of matrix multiplication. We can also apply the promotion to pointwise multiplication $\times$.

We will exploit the following property:

$$r(M \times c) = (r \times c^T)M, \qquad r \text{ can be a matrix} \tag{1}$$

$$(r \times M)c = M(c \times r^T), \qquad c \text{ can be a matrix.} \tag{2}$$

Note that $r \times c^T$ gives a row vector, while $r \times c$ would give a matrix (both r and c would be promoted first).

2. Initial computations

We start with

$$\Gamma(L, f, h) = \frac{1}{2}\left(L(f \times h) - f \times Lh - h \times Lf\right),$$

which is a column vector. We are interested in

$$\begin{aligned}\Gamma_2(L, f, f) &= \frac{1}{2}L\Gamma(L, f, f) - \Gamma(L, Lf, f) = \cdots = \\ &= \frac{1}{4}L^2 f^{\times 2} - L(f \times Lf) + \frac{1}{2}f \times L^2 f + \frac{1}{2}(Lf)^{\times 2}. \end{aligned} \tag{3}$$

We have a squared distance matrix D (size $n \times x$), Laplace matrix L, and a set of functions f, one for each entry of D (pair of points).

We have $f^{(i,j)} = D_{:i} - D_{:,j}$. We take a difference of two columns. Hence we have n^2 functions with n entries each. This makes it hard to store all of them.

Each function $f^{(i,j)}$ is used to find one entry $R_{i,j}$ in the Ricci matrix:

$$R_{i,j} = \Gamma_2(L, f^{(i,j)}, f^{(i,j)})_i.$$

We take i-th row from a column vector $\Gamma_2(\ldots)$.

3. One row in Ricci matrix

We will try to compute the whole row of the Ricci matrix at once.

We define the functions matrix $F(i) = D_{:i} - D$. Then $F(i)$ encodes all functions $f^{(i,\cdot)}$ as its columns.

We can use $F(i)$ in place of f in the formula (3) for Γ_2. We get

$$R_i = \left(\frac{1}{4}L^2F(i)^{\times 2} - L(F(i) \times LF(i)) + \frac{1}{2}F(i) \times L^2F(i) + \frac{1}{2}(LF(i))^{\times 2}\right)_i.$$

Since we are computing just one row, we can do that in each term. To further simplify the notation we drop (i) from F.

Note that for matrices $(A \times B)_i = A_i \times B_i$ and $(AB)_i = A_iB$. Hence

$$R_i = \frac{1}{4}(L^2)_iF^{\times 2} - L_i(F \times LF) + \frac{1}{2}F_i \times (L^2)_iF + \frac{1}{2}(LF)_i^{\times 2}. \tag{4}$$

First note that

$$LF = L(D_{:i} - D) = (LD)_{:i} - LD.$$

We can precompute a few matrix products independent of i:

$$\begin{aligned} A &= L^2, \\ B &= LD, \\ C &= AD. \end{aligned}$$

Note that B and C are not symmetric, since products of symmetric matrices generally are not symmetric.

Now we will look at each term in (4). The last term is the simplest:

$$\begin{aligned} (LF)_i^{\times 2} = (B_{:i} - B)_i^{\times 2} &= (B_{ii} - B_i)^{\times 2} \\ &= (B^{\times 2})_i + (B^{\times 2})_{ii} - 2B_i \times B_{ii}. \end{aligned}$$

Hence it can be computed as row$\times$row operation.

The first term can also be simplified:

$$\begin{aligned} (L^2)_iF^{\times 2} &= A_i(D^{\times 2} + D_{:i}^{\times 2} - 2D \times D_{:i}) \\ &= A_iD^{\times 2} + A_i(D^{\times 2})_{:i} - 2A_i(D \times D_{:i}) \\ &= (AD^{\times 2})_i + (AD^{\times 2})_{ii} - 2(A \times D)_iD, \end{aligned}$$

In the last step we used symmetry of D and (1). For nonsymmetric matrix column $D_{:,i}$ would simply need to be transposed instead of changing to D_i. Next we have

$$\begin{aligned}
F_i \times (L^2)_i F &= (D_{ii} - D_i) \times [A_i(D_{:i} - D)] \\
&= (D_{ii} - D_i) \times ((AD)_{ii} - (AD)_i) \\
&= (D_{ii} - D_i) \times (C_{ii} - C_i) \\
&= (D \times C)_i + (D \times C)_{ii} - D_i \times C_{ii},
\end{aligned}$$

since $D_{ii} = 0$. Finally

$$\begin{aligned}
L_i(F \times LF) &= L_i((D_{:i} - D) \times (B_{:i} - B)) = \\
&= L_i \left(D \times B + (D \times B)_{:i} - D \times B_{:i} - B \times D_{:i} \right) = \\
&= G_i + G_{ii} - (L_i \times (B^T)_i)D - (L_i \times D_i)B \\
&= G_i + G_{ii} - (L \times B^T)_i D - (L \times D)_i B,
\end{aligned}$$

where $G = L(D \times B)$.

To make our notation even more compact we let $diag(M)$ be a column vector composed of the diagonal elements of M. Then

$$\overline{M} = M + diag(M)$$

satisfies

$$\overline{M}_i = M_i + M_{ii}.$$

With this notation we can represent a row of Ricci matrix (given by (4)) as

$$\begin{aligned}
R_i = &\frac{1}{4}\left(\overline{AD^{\times 2}}\right)_i - \frac{1}{2}(A \times D)_i D \\
&- \left(\overline{L(D \times B)}\right)_i + (L \times B^T)_i D + (L \times D)_i B \\
&+ \frac{1}{2}\left(\overline{D \times C}\right)_i - \frac{1}{2} D_i \times C_{ii} \\
&+ \frac{1}{2}\left(\overline{B^{\times 2}}\right)_i - B_i \times B_{ii}.
\end{aligned}$$

4. Full Ricci matrix

Note that the formula for the row of the Ricci matrix contains only take a row operations for various matrices. We can also combine all overline operations. Hence

$$
\begin{aligned}
R = &\overline{\frac{1}{4}AD^{\times 2} - L(D \times B) + \frac{1}{2}D \times C + \frac{1}{2}B^{\times 2}} \\
&+ \left((L \times B^T) - \frac{1}{2}(A \times D)\right) D + (L \times D)B \\
&- \frac{1}{2}D \times diag(C) - B \times diag(B),
\end{aligned} \tag{5}
$$

with $A = L^2$, $B = LD$ and $C = AD$, a *diag* denoting a column vector composed of the diagonal of a matrix.

5. Order of computations

We want to minimize the number of multiplications while keeping memory requirement small. Clearly, 7 matrix multiplications and 3 extra matrices are enough. Although it should be possible to bring this down to 2 extra matrices. Furthermore, Laplacian could be just computed as needed, since it is really fast.

5.1. **Only 2 temporary matrices.** We will avoid computing A by taking L out of the first two terms.

- Compute $B \leftarrow D^{\times 2}/4$ first, then $A \leftarrow LB$ (1st M-M multiply).
- Compute $B \leftarrow LD$ (2nd M-M multiply), then $A \leftarrow A - D \times B$.
- Now $R \leftarrow LA$ (3rd M-M multiply).
- Compute $A \leftarrow LB$ (4th M-M multiply, called C in the section above), then finish adding *diag* terms and overline.

Now we need the second line terms from (5). We already have B computed.

- Find $A \leftarrow LL$ (5th M-M multiply), then $A \leftarrow \left((L \times B^T) - \frac{1}{2}(A \times D)\right)$.
- Now $R \leftarrow R + AD$ (6th M-M multiply, full dgemm).
- Finally, $A \leftarrow D \times L$, then $R \leftarrow R + AB$ (7th M-M multiply, full dgemm).