(Parallel, Randomized)

Rank-revealing factorizations

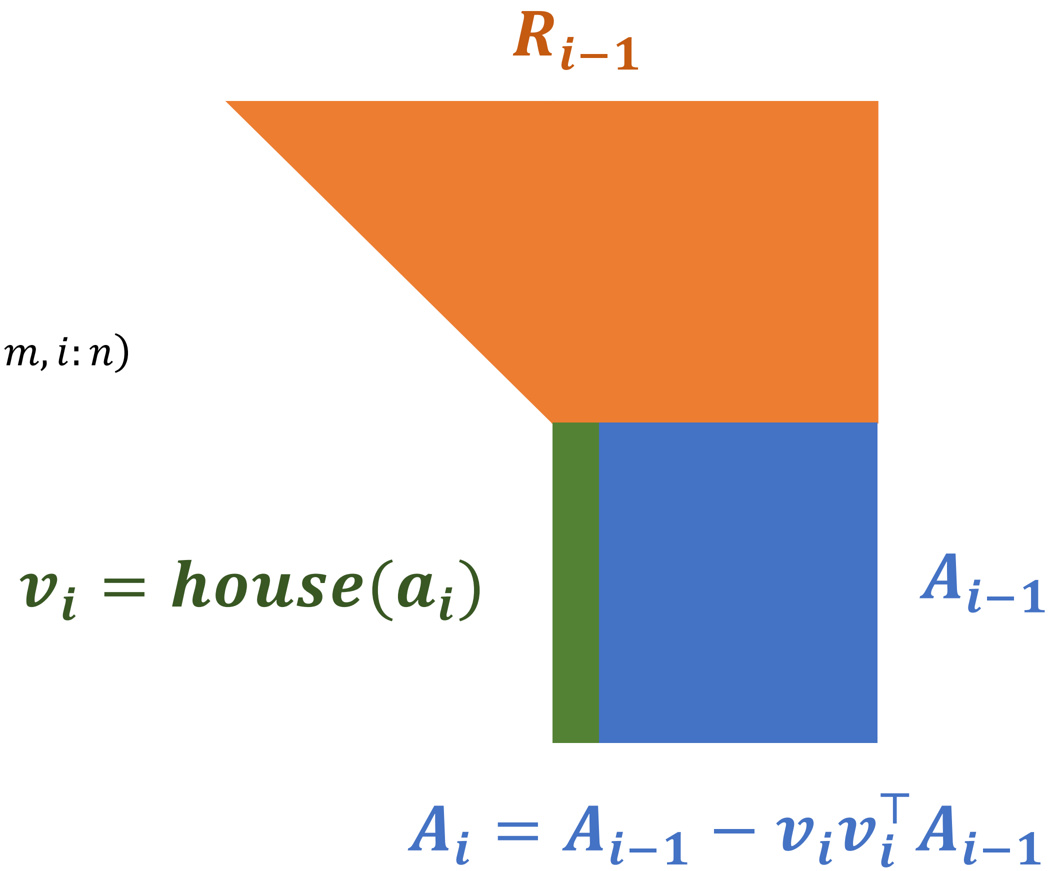
Léopold Cambier

February 2020

QR

for $i = 1 \ldots \min(m, n)$

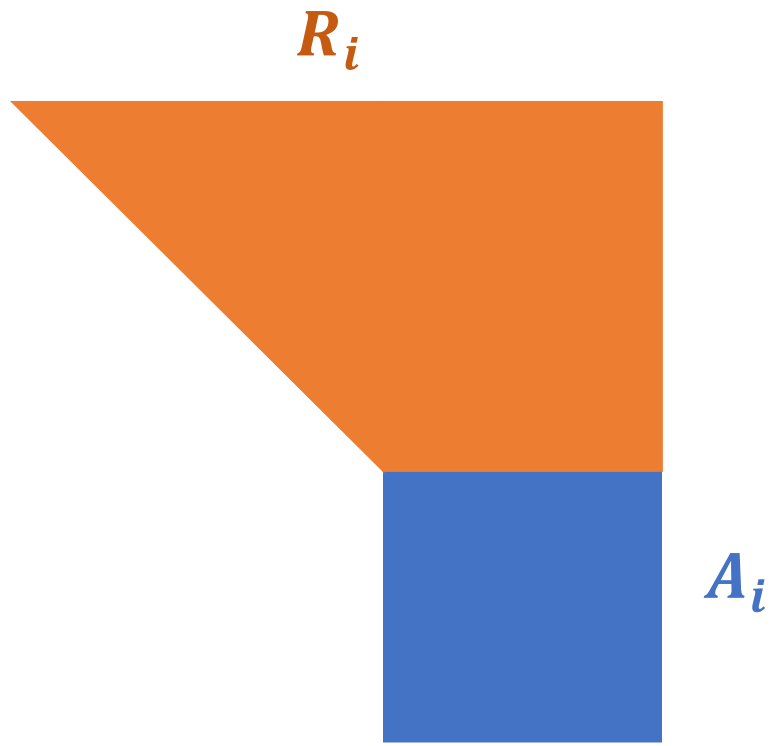
- $v_i = house(A(i:m, i))$
- $A(i:m, i:n) = \left(I - v_i v_i^\top\right) A(i:m, i:n)$



QR

for $i = 1 \ldots \min(m, n)$

- $v_i = house(A(i{:}m, i))$
- $A(i{:}m, i{:}n) = \left(I - v_i v_i^\top\right) A(i{:}m, i{:}n)$



Low-rank approximations

- We know the SVD

$$A = USV \approx U_r \Sigma_r V_r, \quad ||A - U_r \Sigma_r V_r||_2 = \sigma_{r+1}(A)$$

- If we add column pivoting to QR, we can do "almost" the same

$$AP = QR$$

$$Q^\top AP = \begin{bmatrix} R_{11} & R_{12} \\ & R_{22} \end{bmatrix}$$

If $||R_{22}||_2$ is small then

Low-rank

$$A \approx Q_1 [R_{11} \quad R_{12}] P^\top = \boxed{Q_1 W}$$

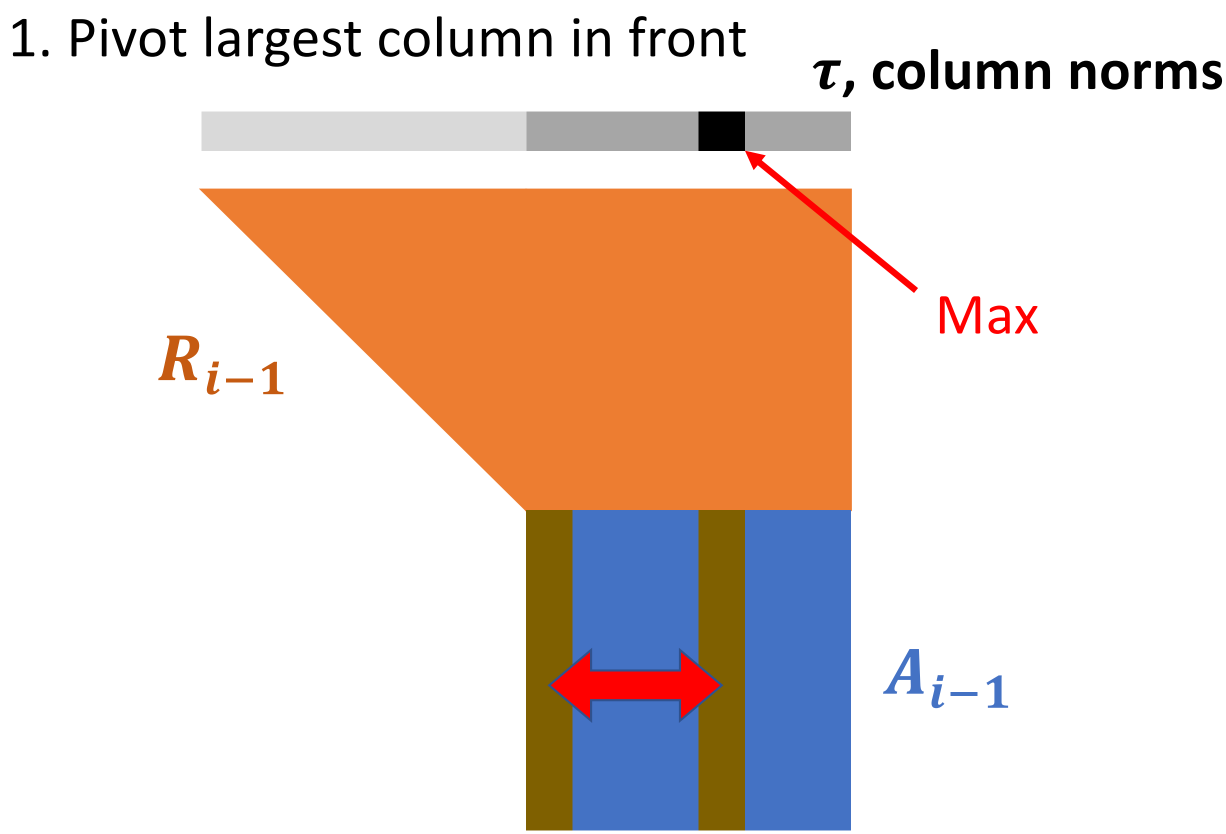
(+) Cheaper than SVD (direct - not "iterative")

(-) Less reliable

(+) In practice good enough

QRCP

"Classical algorithm"
(Golub & Van Loan,
Algorithm 5.4.1)

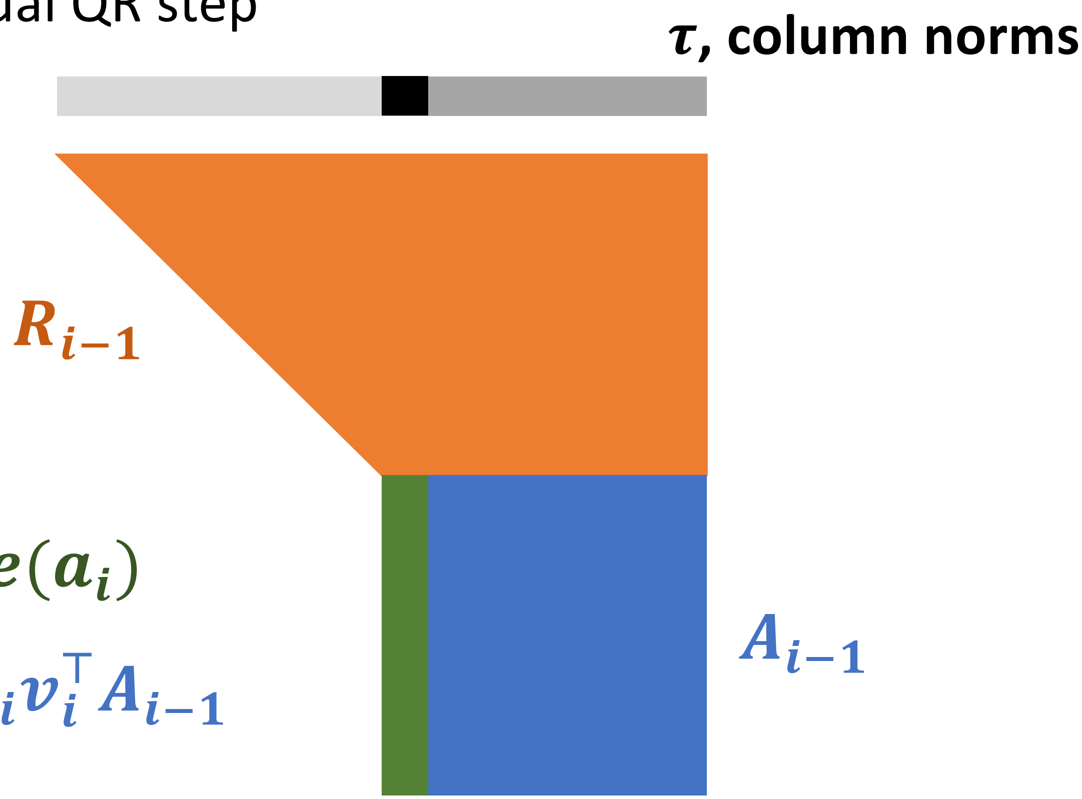


Golub, Gene H., and Charles F. Van Loan.
Matrix computations. 4th edition. JHU press, 2013.

QRCP

"Classical algorithm" (Golub & Van Loan, Algorithm 5.4.1)

2. Usual QR step



$$v_i = house(a_i)$$

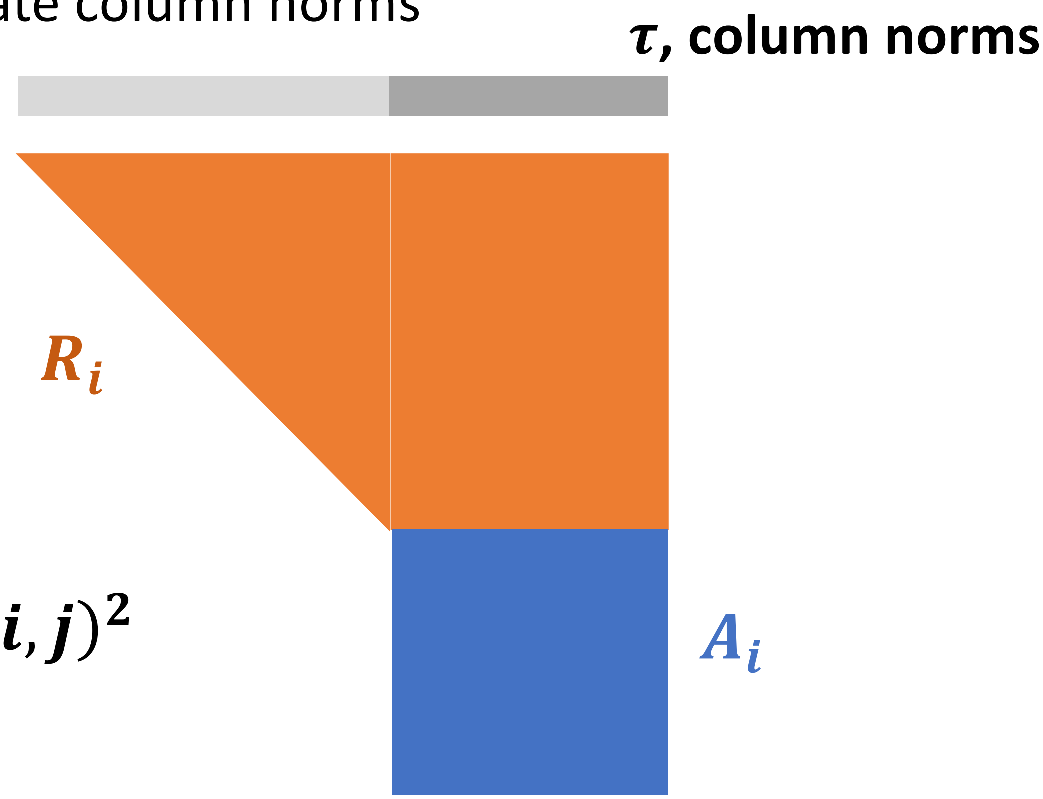
$$A_i = A_{i-1} - v_i v_i^\top A_{i-1}$$

Golub, Gene H., and Charles F. Van Loan. *Matrix computations*. 4th edition. JHU press, 2013.

QRCP

"Classical algorithm" (Golub & Van Loan, Algorithm 5.4.1)

3. Update column norms



$$\tau_j = \tau_j - A_i(i,j)^2$$

Golub, Gene H., and Charles F. Van Loan.
Matrix computations. 4th edition. JHU press, 2013.

QRCP

geqpf in Lapack

```
for j = 1:n
    c(j) = A(1:m, j)^T A(1:m, j)
end
r = 0
τ = max{c(1), ..., c(n)}
while τ > 0 and r < n
    r = r + 1
    Find smallest k with r ≤ k ≤ n so c(k) = τ.
    piv(r) = k
    A(1:m, r) ↔ A(1:m, k)
    c(r) ↔ c(k)
    [v, β] = house(A(r:m, r))
    A(r:m, r:n) = (I_{m-r+1} − βvv^T)A(:r:m, r:n)
    A(r + 1:m, r) = v(2:m − r + 1)
    for i = r + 1:n
        c(i) = c(i) − A(r, i)^2
    end
    τ = max{c(r + 1), ..., c(n)}
end
```

Pivot max in front

Usual QR step

Update norms

Golub, Gene H., and Charles F. Van Loan.
Matrix computations. 4th edition. JHU press, 2013.

QRCP

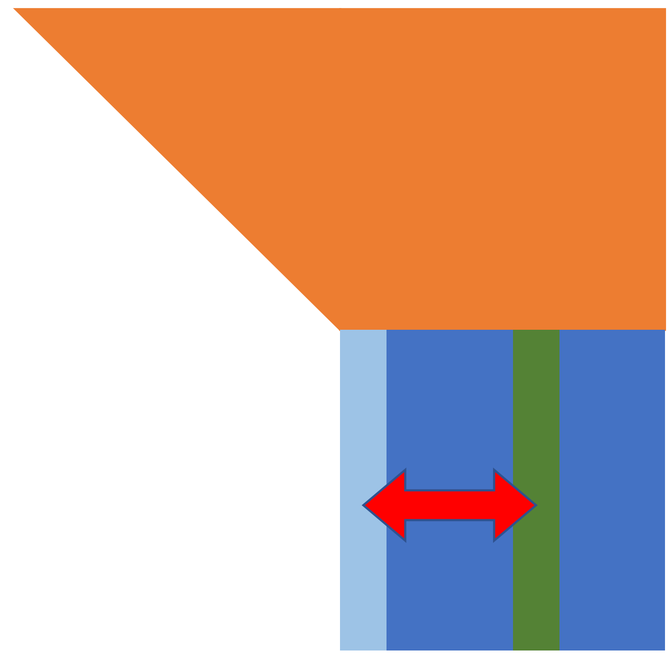
- Very reliable in practice to reveal the rank (not guaranteed!)
- Previous algorithm not blocked (BLAS2 - geqpf)
- In practice algorithm can be blocked (BLAS3 - geqp3)

- Issue:

1. **Very sequential: Need step < j for step j.**
2. **Small "blocks": 1 column at a time.**
3. **Pivoting: Hard in parallel.**

Column pivoting ≈ Range finding

Pivoting ≈ range approximation
(of trailing matrix)



Randomized range approximation

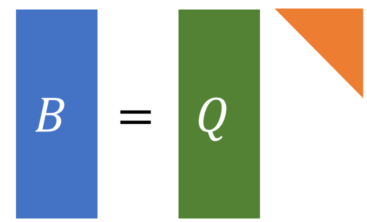
Ω i.i.d. Gaussian, size $n \times b$

$$B = A\Omega$$

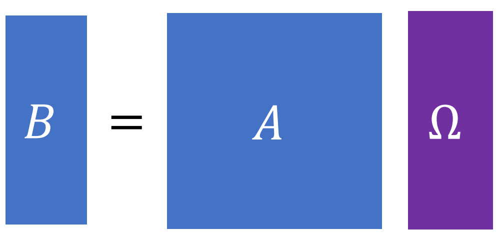
$Q = \mathrm{qr}(B)$ (**random**), orthogonal s.t.

$$\mathrm{range}(A) \approx \mathrm{range}(Q)$$

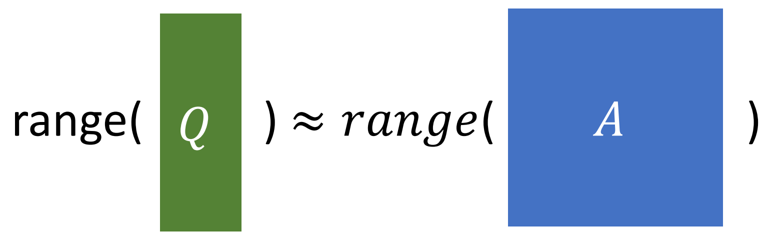
2.



1.



3.



Halko, Nathan, Per-Gunnar Martinsson, and Joel A. Tropp.
"Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions." *SIAM review* 53.2 (2011): 217-288.

Randomization is "good enough"

Expectation

THEOREM 1.1. *Suppose that $\boldsymbol{A}$ is a real $m \times n$ matrix. Select a target rank $k \geq 2$ and an oversampling parameter $p \geq 2$, where $k + p \leq \min\{m, n\}$. Execute the proto-algorithm with a standard Gaussian test matrix to obtain an $m \times (k + p)$ matrix $\boldsymbol{Q}$ with orthonormal columns. Then*

Small poly(k), decay w/ p

$$\mathbb{E}\,\|\boldsymbol{A} - \boldsymbol{Q}\boldsymbol{Q}^*\boldsymbol{A}\| \leq \left[1 + \frac{4\sqrt{k+p}}{p-1} \cdot \sqrt{\min\{m,n\}}\right] \sigma_{k+1},$$

"Spectral" accuracy

where $\mathbb{E}$ denotes expectation with respect to the random test matrix and σ_{k+1} is the $(k+1)$th singular value of $\boldsymbol{A}$.

With high probability

As we discuss in §10.3, the probability that the error satisfies

$$\|\boldsymbol{A} - \boldsymbol{Q}\boldsymbol{Q}^*\boldsymbol{A}\| \leq \left[1 + 11\sqrt{k+p} \cdot \sqrt{\min\{m,n\}}\right] \sigma_{k+1}$$

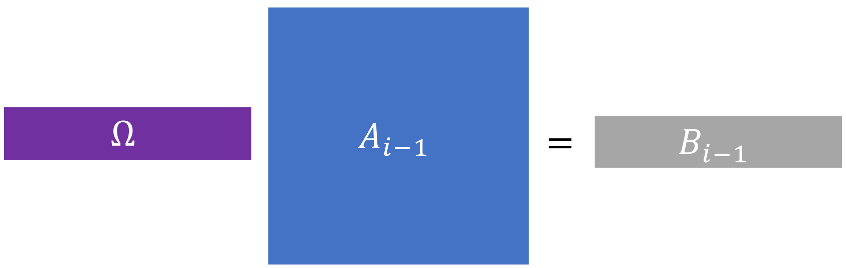
is at least $1 - 6 \cdot p^{-p}$ under very mild assumptions on p.

High probability w/ p

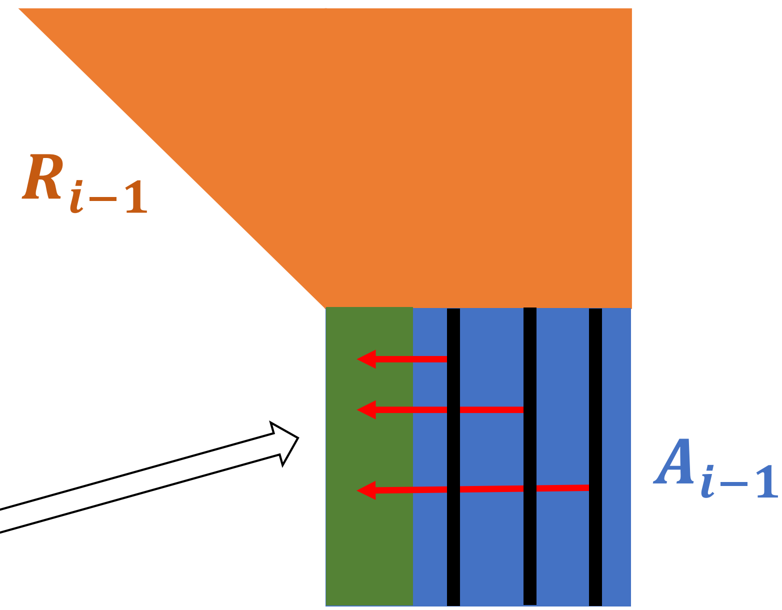
Halko, Nathan, Per-Gunnar Martinsson, and Joel A. Tropp.
"Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions." *SIAM review* 53.2 (2011): 217-288.

(1) QR w/ randomized block pivoting

1. Compute Ω random and

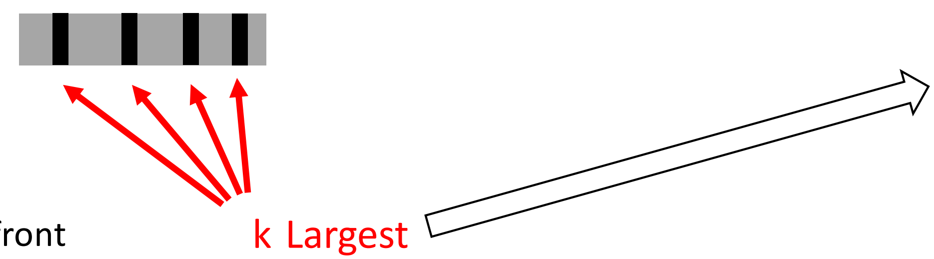


3. Block QR step



2. Pick a set of pivots by running QRCP on B_{i-1}

$$B_{i-1}P = Q_i R_i$$

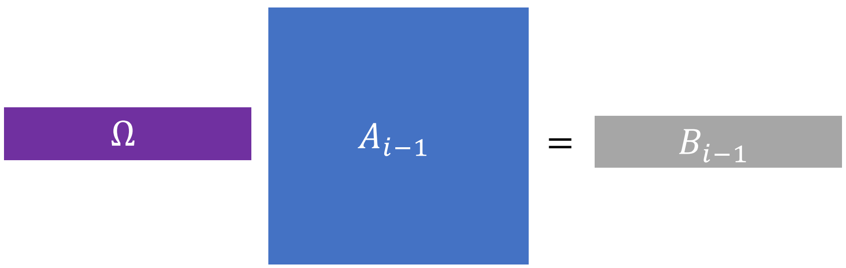


Bring first ***k*** columns in front

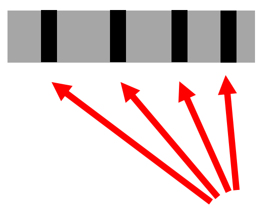
Xiao, Jianwei, Ming Gu, and Julien Langou. "Fast parallel randomized QR with column pivoting algorithms for reliable low-rank matrix approximations." *2017 IEEE 24th International Conference on High Performance Computing (HiPC)*. IEEE, 2017.

(1) QR w/ randomized block pivoting

1. Compute Ω random and



2. Pick a set of pivot by running QRCP on B_{i-1}



$$B_{i-1}P = Q_i R_i$$

Bring first $\boldsymbol{k}$ columns in front

k Largest

3. Block QR step

Still lots of pivoting and shuffle. Hard to parallelize.

Xiao, Jianwei, Ming Gu, and Julien Langou. "Fast parallel randomized QR with column pivoting algorithms for reliable low-rank matrix approximations." *2017 IEEE 24th International Conference on High Performance Computing (HiPC)*. IEEE, 2017.

(2) Randomized range approximation

We don't really care about R…

Build $n \times k$ Gaussian Ω

Form $Y = A\Omega$

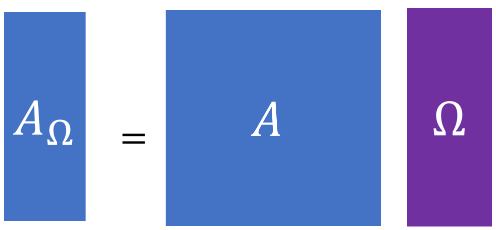
Compute $Q = qr(Y)$

Set $A \approx QW$ with $W = Q^\top A$

Low-rank

(2) Blocked, adaptive, randomized range approx.

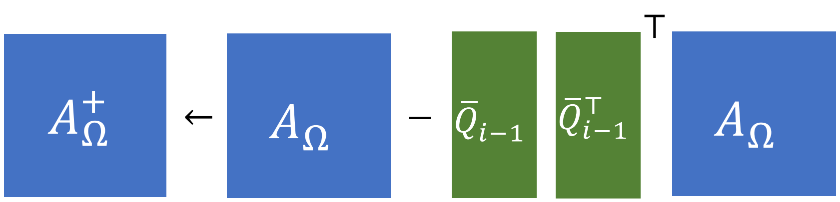
1. Compute $A_\Omega = A\Omega$



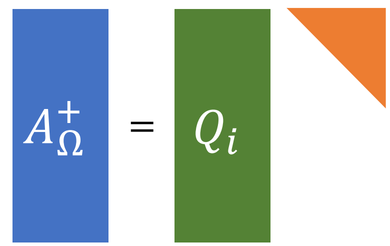
$$\text{range}(A_\Omega) \approx \text{range}(A)$$

2. Project* out <u>all</u> previous Q_i's

$$A_\Omega^+ \leftarrow A_\Omega - \bar{Q}_{i-1}\bar{Q}_{i-1}^\top A_\Omega$$



3. QR on A_Ω



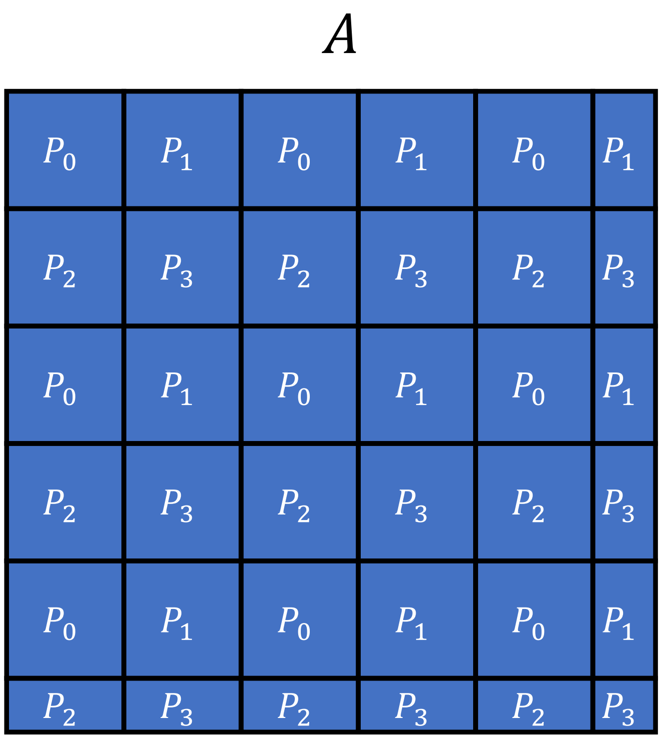
$$\text{range}(Q_i) = \text{range}(A_\Omega^+)$$
$$= range\left(\left(I - \bar{Q}_{i-1}\bar{Q}_{i-1}^\top\right)A_\Omega\right)$$
$$\approx range\left(\left(I - \bar{Q}_{i-1}\bar{Q}_{i-1}^\top\right)A\right)$$

4. Repeat

$$\bar{Q}_i = Q_i\bar{Q}_{i-1}$$

* Stability ! HH or MGS.

(2) Blocked, adaptive, randomized range approx.



Main operations:

- $B = A\Omega$ — Matrix-matrix. Very parallel.
- $B = QR$ — Block QR, cheap if k not too large
- $A^{+} = A - QQ^{\top}A$ — Sequential per column (HH)
  All columns independent

References

“Classical” algorithm

- Golub, Gene H., and Charles F. Van Loan. *Matrix computations*. 4th edition. JHU press, 2013.

Block QRCP

- Quintana-Ortí, Gregorio, Xiaobai Sun, and Christian H. Bischof. "A BLAS-3 version of the QR factorization with column pivoting." *SIAM Journal on Scientific Computing* 19.5 (1998): 1486-1494.

Parallelization

- Tomás, Andrés, Zhaojun Bai, and Vicente Hernández. "Parallelization of the QR decomposition with column pivoting using column cyclic distribution on multicore and GPU processors." *International Conference on High Performance Computing for Computational Science*. Springer, Berlin, Heidelberg, 2012.

References

Improved column pivoting

- Xiao, Jianwei, Ming Gu, and Julien Langou. "Fast parallel randomized QR with column pivoting algorithms for reliable low-rank matrix approximations." *2017 IEEE 24th International Conference on High Performance Computing (HiPC)*. IEEE, 2017.
- Demmel, James W., et al. "Communication avoiding rank revealing QR factorization with column pivoting." *SIAM Journal on Matrix Analysis and Applications* 36.1 (2015): 55-89.
- Martinsson, Per-Gunnar, et al. "Householder QR factorization with randomization for column pivoting (HQRRP)." *SIAM Journal on Scientific Computing* 39.2 (2017): C96-C115.

References

Range approximation

- Halko, Nathan, Per-Gunnar Martinsson, and Joel A. Tropp. "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions." *SIAM review* 53.2 (2011): 217-288.
- Martinsson, Per-Gunnar. "Randomized methods for matrix computations." *The Mathematics of Data* 25 (2019): 187-231.