The Generalized Matrix Chain Algorithm

Henrik Barthels, Marcin Copik, Paolo Bientinesi

HPAC High Performance and Automatic Computing
RWTH AACHEN UNIVERSITY

The Linear Algebra Mapping Problem

- How to compute the following expressions?

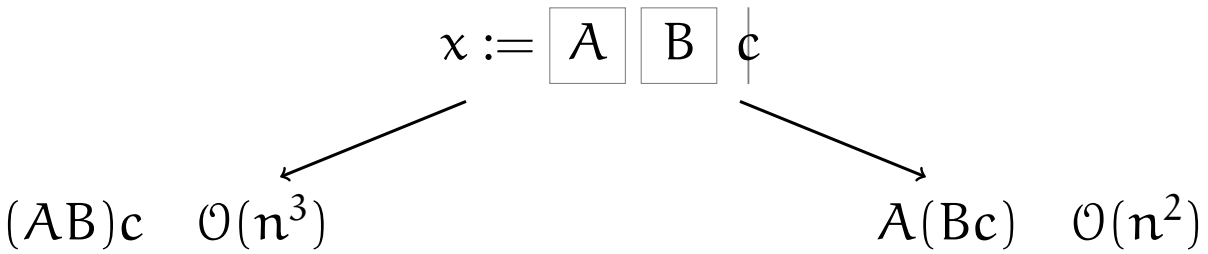
$$b := (X^T X)^{-1} X^T y$$
$$b := (X^T M^{-1} X)^{-1} X^T M^{-1} y$$
$$x := W(A^T (AWA^T)^{-1} b - c)$$
$$x := \left(A^{-T} B^T B A^{-1} + R^T[\Lambda(Rz)]R\right)^{-1} A^{-T} B^T B A^{-1} y$$

- High-level languages are easy to use, but performance is usually suboptimal.
- BLAS and LAPACK can be fast, but require a lot of expertise.

BLAS [DDC+90], **LAPACK** [AB+99]

- $y \leftarrow Ax + y$
- $C \leftarrow AB + C$
- $B \leftarrow A^{-1} B$
- $\dots$

The Matrix Chain Problem



- Find optimal parenthesization of $A_1 A_2 \cdots A_k$.
- Dynamic programming $\mathcal{O}(k^3)$ algorithm [CRL90].

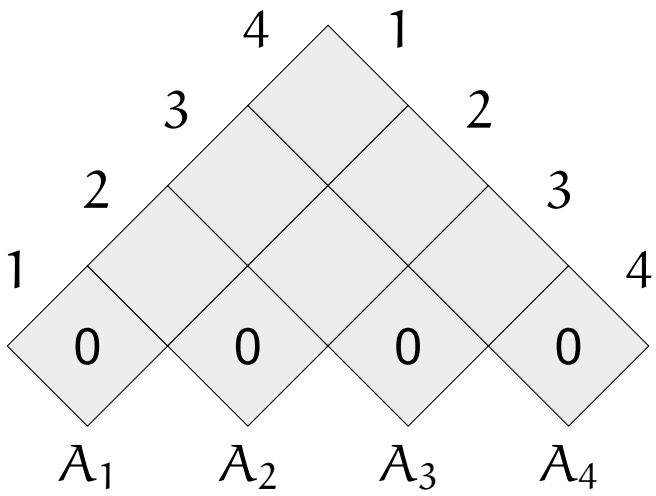
The Generalized Matrix Chain Problem

$$A^T B C^{-1} D$$

- Unary operators: Transposition, inversion ($\Rightarrow$ linear systems).
- Matrix properties: Upper/lower triangular, symmetric, SPD, diagonal,…

The Matrix Chain Algorithm

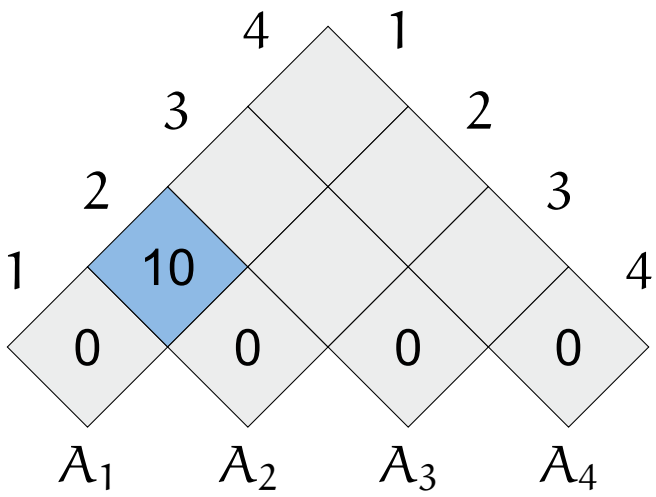
Example: $A_1A_2A_3A_4$



Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$

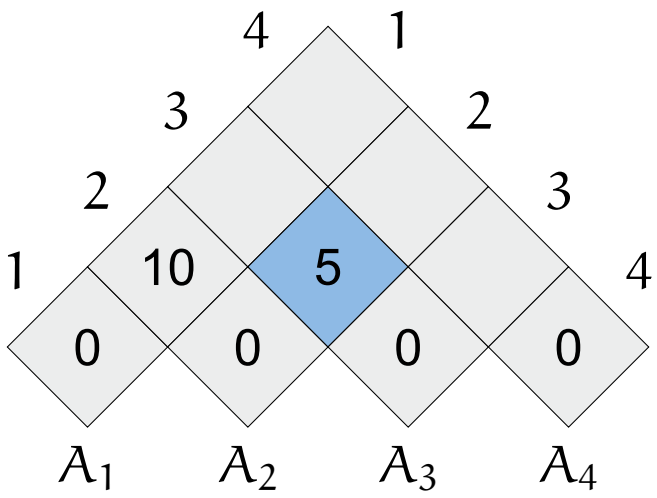


$$c(A_1 \cdot A_2) = 10$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$

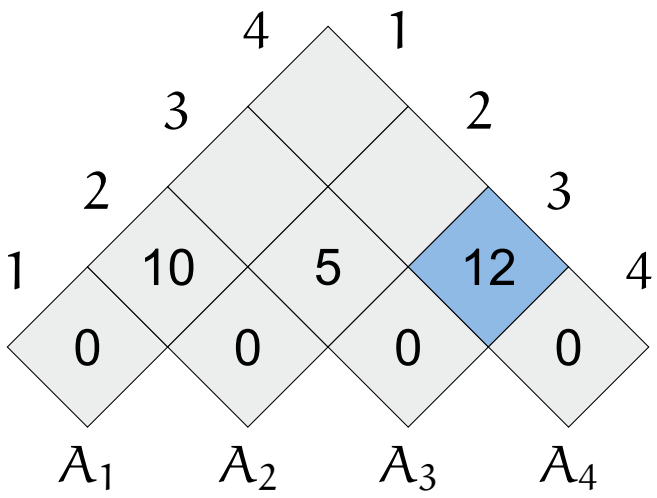


$$c(A_2 \cdot A_3) = 5$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$

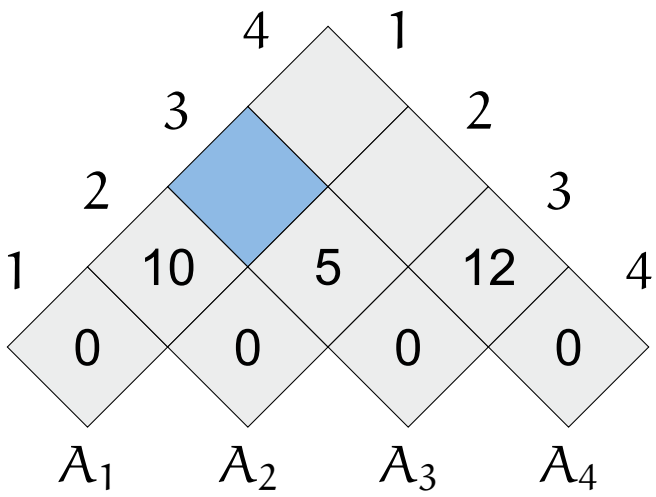


$$c(A_3 \cdot A_4) = 7$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$

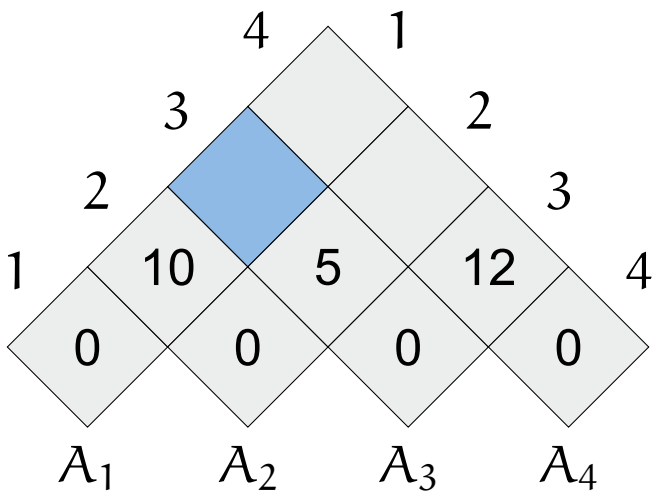


How to parenthesize $A_1A_2A_3$?

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$

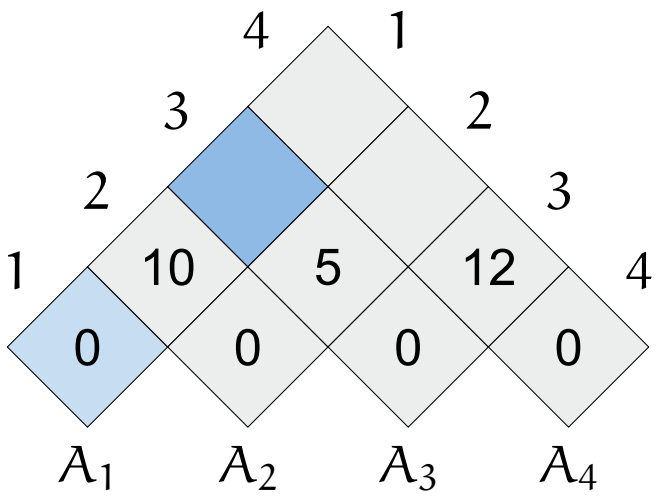


How to parenthesize $A_1A_2A_3$?
- $A_1(A_2A_3)$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



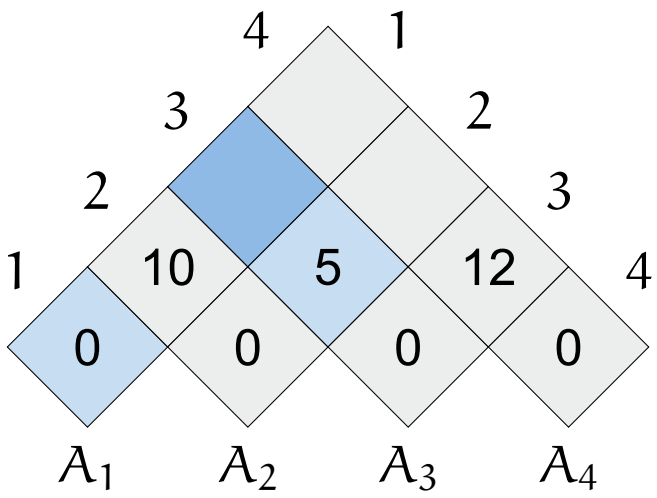
How to parenthesize $A_1A_2A_3$?
- $A_1(A_2A_3)$

$$c(A_1) = 0$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



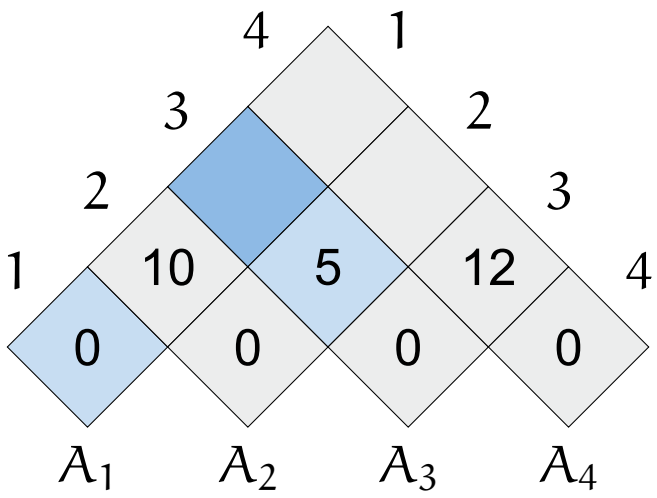
How to parenthesize $A_1A_2A_3$?
- $A_1(A_2A_3)$

$$c(A_1) = 0$$
$$c(A_2A_3) = 5$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



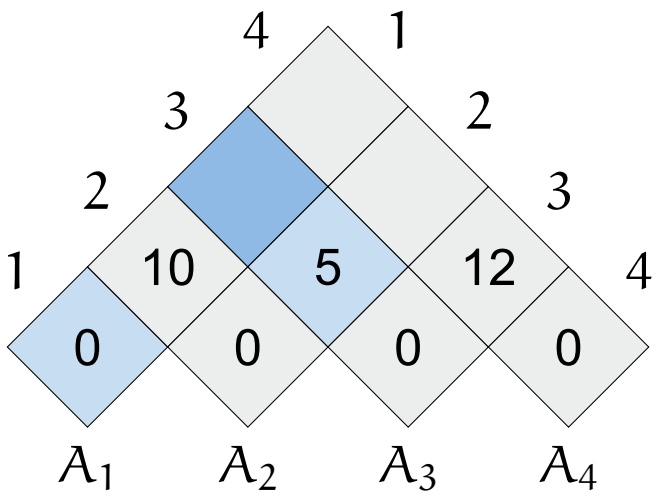
How to parenthesize $A_1A_2A_3$?
- $A_1(A_2A_3)$

$$c(A_1) = 0$$
$$c(A_2A_3) = 5$$
$$c(A_1(A_2A_3)) = 8$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



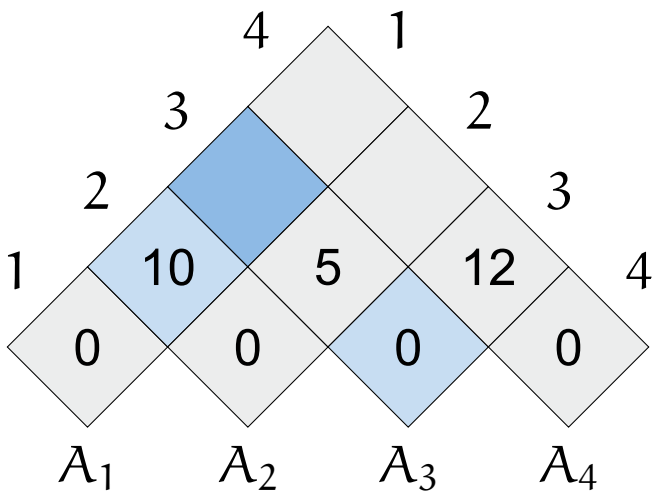
How to parenthesize $A_1A_2A_3$?
- $A_1(A_2A_3)$

$$c(A_1) = 0$$
$$c(A_2A_3) = 5$$
$$c(A_1(A_2A_3)) = 8$$
$$0 + 5 + 8 = 13$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



How to parenthesize $A_1A_2A_3$?

- $A_1(A_2A_3)$

$$c(A_1) = 0$$
$$c(A_2A_3) = 5$$
$$c(A_1(A_2A_3)) = 8$$
$$0 + 5 + 8 = 13$$

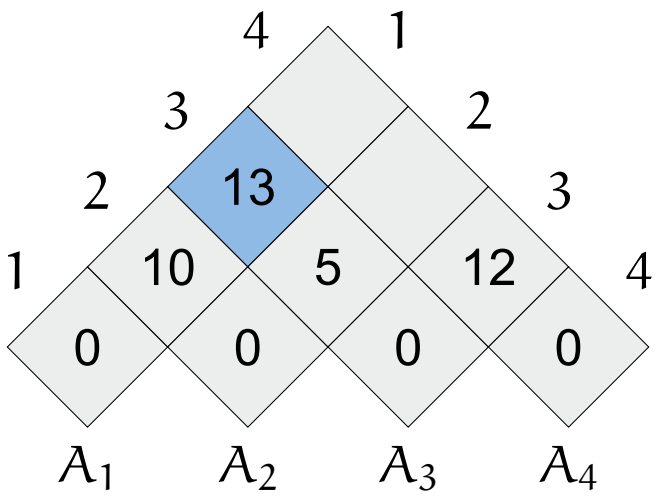
- $(A_1A_2)A_3$

$$10 + 0 + 6 = 16$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



How to parenthesize $A_1A_2A_3$?

- $A_1(A_2A_3)$

$$c(A_1) = 0$$
$$c(A_2A_3) = 5$$
$$c(A_1(A_2A_3)) = 8$$
$$0 + 5 + 8 = 13$$

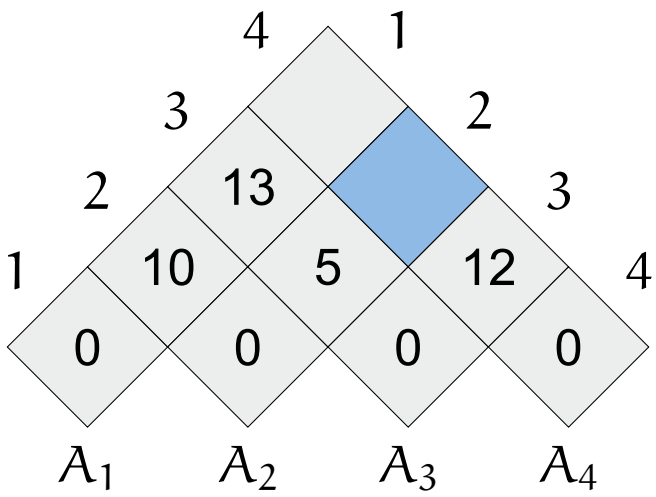
- $(A_1A_2)A_3$

$$10 + 0 + 6 = 16$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$

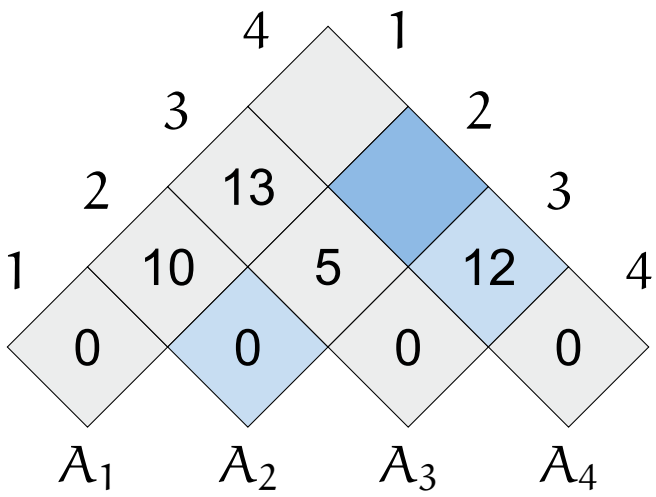


How to parenthesize $A_2A_3A_4$?

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$

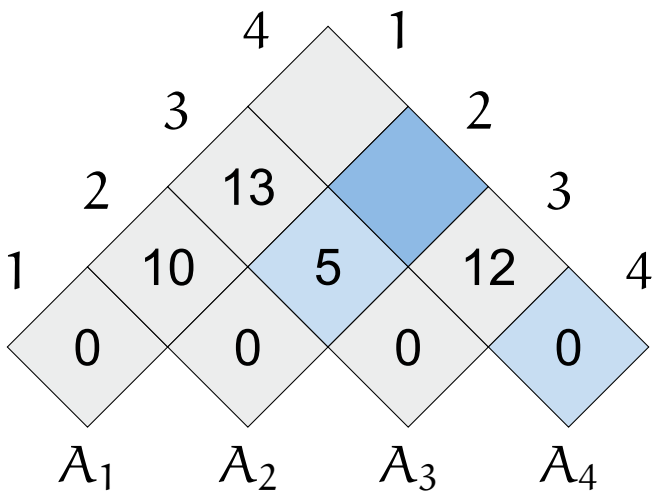


How to parenthesize $A_2A_3A_4$?
- $A_2(A_3A_4)$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



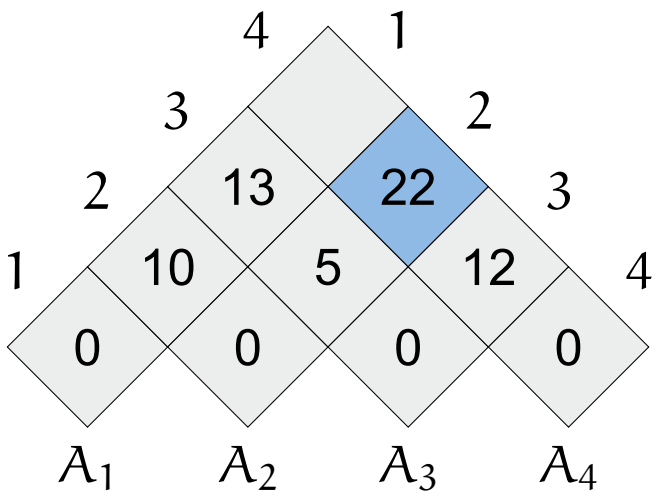
How to parenthesize $A_2A_3A_4$?

- $A_2(A_3A_4)$
- $(A_2A_3)A_4$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



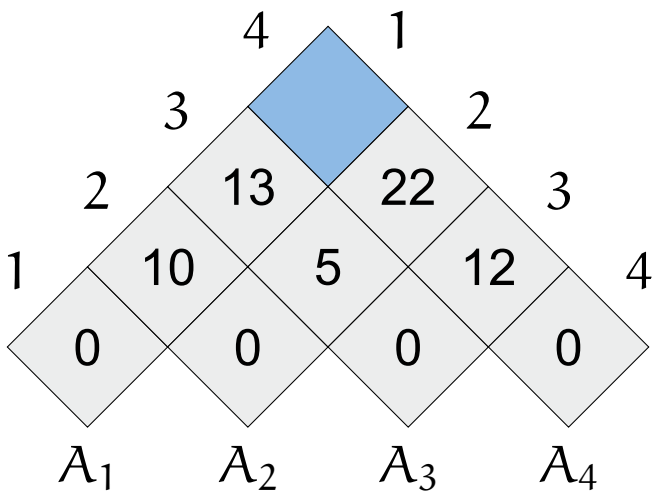
How to parenthesize $A_2A_3A_4$?

- $A_2(A_3A_4)$
- $(A_2A_3)A_4$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$

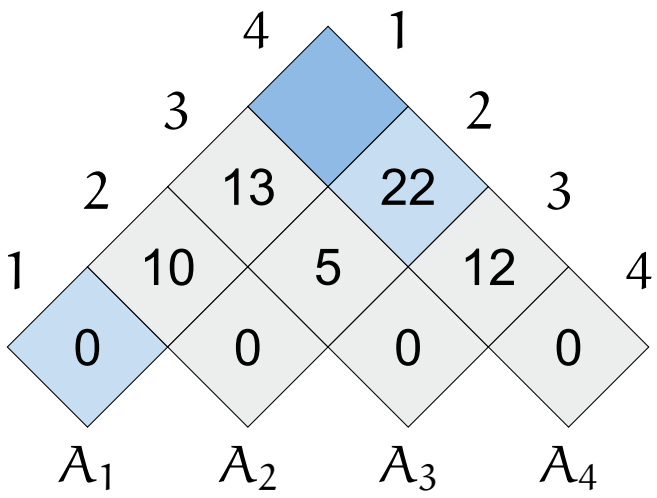


How to parenthesize $A_1A_2A_3A_4$?

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$

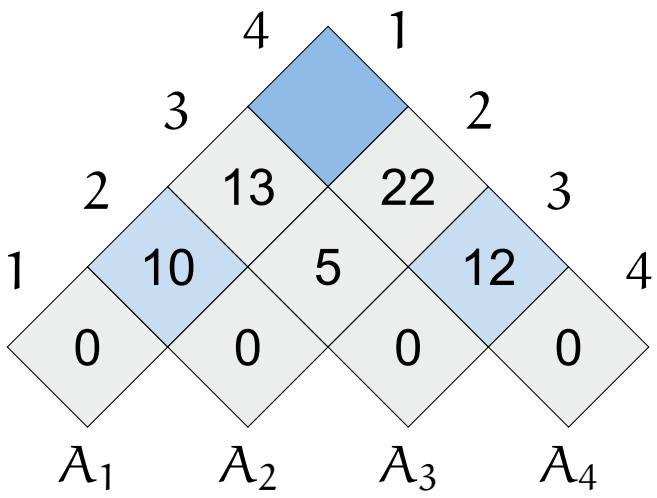


How to parenthesize $A_1A_2A_3A_4$?
- $A_1(A_2A_3A_4)$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



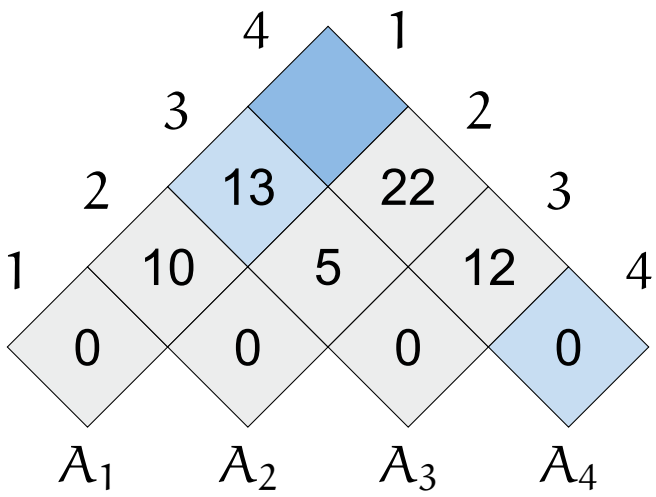
How to parenthesize $A_1A_2A_3A_4$?

- $A_1(A_2A_3A_4)$
- $(A_1A_2)(A_3A_4)$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



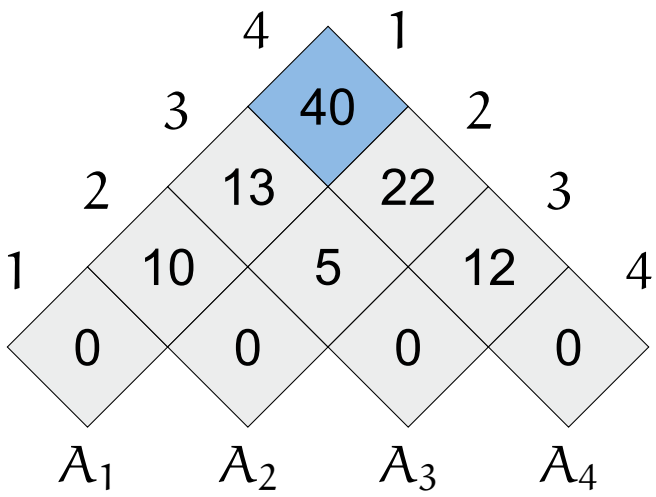
How to parenthesize $A_1A_2A_3A_4$?

- $A_1(A_2A_3A_4)$
- $(A_1A_2)(A_3A_4)$
- $(A_1A_2A_3)A_4$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Matrix Chain Algorithm

Example: $A_1A_2A_3A_4$



How to parenthesize $A_1A_2A_3A_4$?

- $A_1(A_2A_3A_4)$
- $(A_1A_2)(A_3A_4)$
- $(A_1A_2A_3)A_4$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

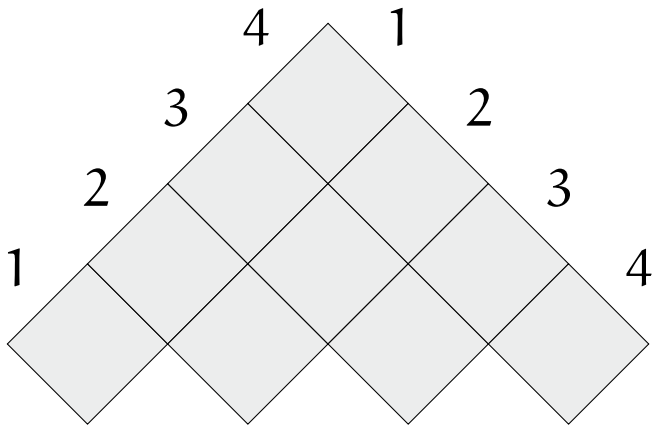
The Generalized Matrix Chain Algorithm

Binary + Unary Operations = Extended Set of Binary Operations

- Multiplication with transposition: AB^T, A^TB, A^TB^T.
- Linear system: AB^{-1}, $A^{-1}B$.
- Linear system with transposition: A^TB^{-1}, AB^{-T}, …
- Special case: $A^{-1}B^{-1}$.

The Generalized Matrix Chain Algorithm

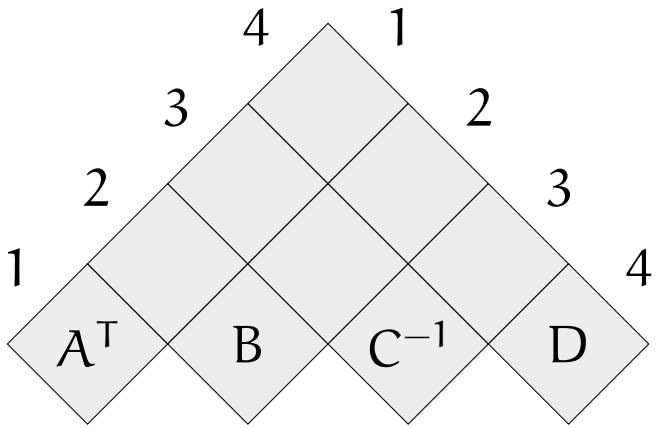
Example: $A^T B C^{-1} D$ A is lower triangular.



Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

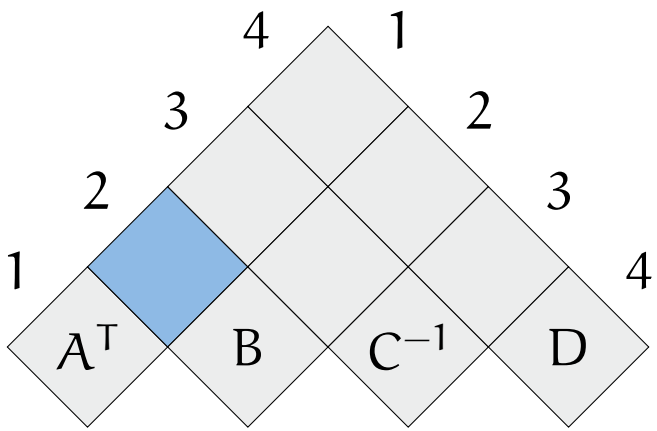
Example: $A^T B C^{-1} D$ A is lower triangular.



Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.



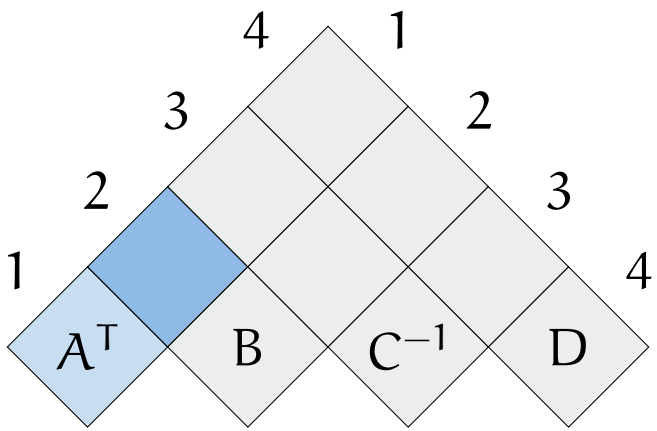
What do we have to compute?

? · ?

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.



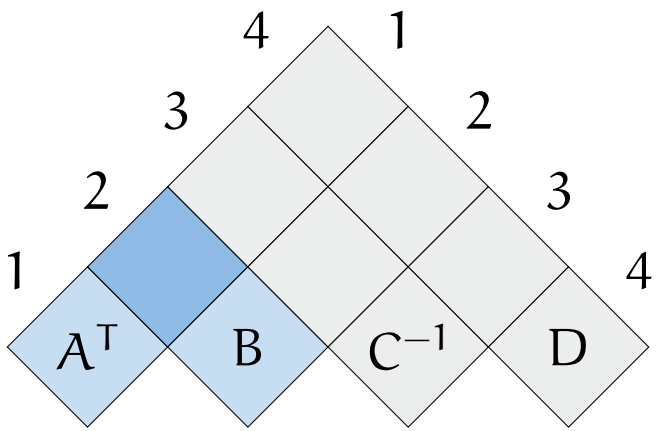
What do we have to compute?

$A^T \cdot ?$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.



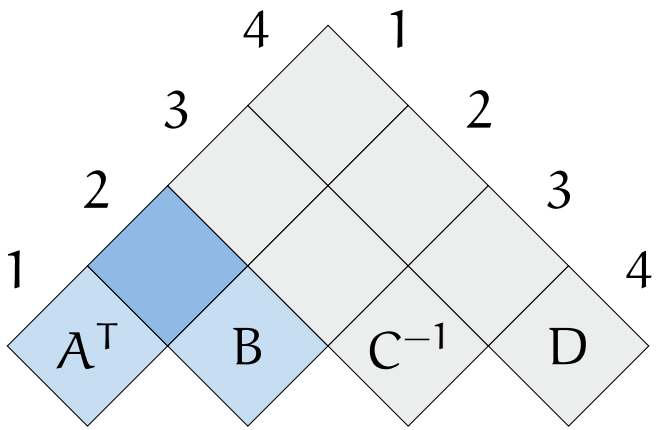
What do we have to compute?

$$A^T \cdot B$$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.



What do we have to compute?

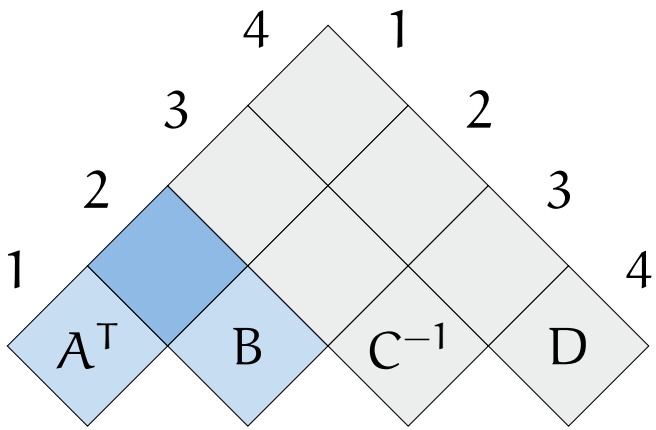
$$A^T \cdot B$$

Name	Pattern	Constraints	Cost [#FLOPs]
GEMM	$X^T Y$	-	$2mnk$
TRMM	$X^T Y$	is_lower_triangular(X)	$m^2 n$
TRMM	$X^T Y$	is_upper_triangular(X)	$m^2 n$
TRSM	$X^{-1} Y$	is_lower_triangular(X)	$m^2 n$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.



What do we have to compute?

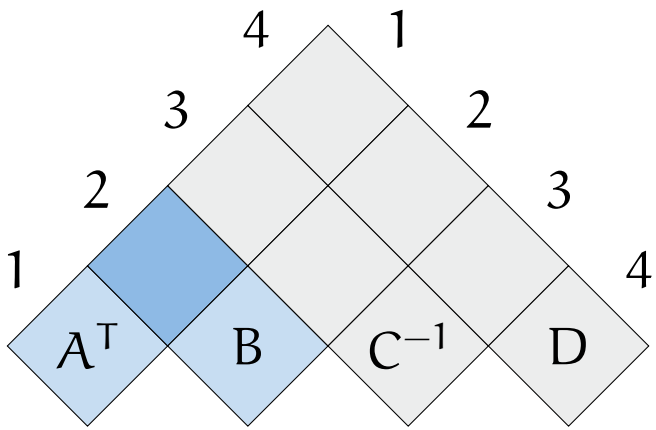
$$A^T \cdot B$$

Name	Pattern	Constraints	Cost [#FLOPs]
GEMM	$X^T Y$	-	$2mnk$
TRMM	$X^T Y$	is_lower_triangular(X)	$m^2 n$
TRMM	$X^T Y$	is_upper_triangular(X)	$m^2 n$
TRSM	$X^{-1} Y$	is_lower_triangular(X)	$m^2 n$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.



What do we have to compute?

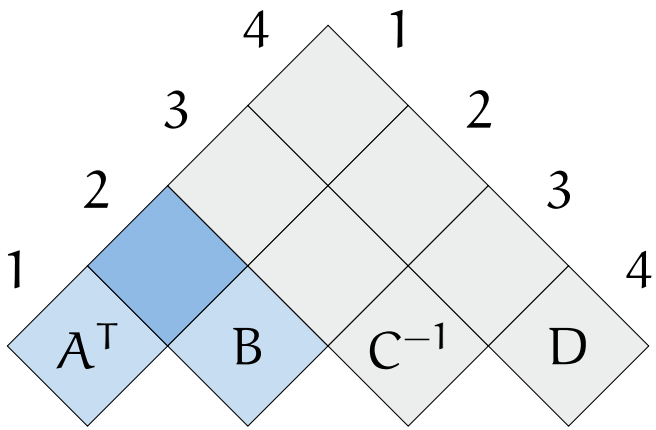
$$A^T \cdot B$$

Name	Pattern	Constraints	Cost [#FLOPs]
GEMM	$X^T Y$	-	$2mnk$
TRMM	$X^T Y$	is_lower_triangular(X)	$m^2 n$
TRMM	$X^T Y$	is_upper_triangular(X)	$m^2 n$
TRSM	$X^{-1} Y$	is_lower_triangular(X)	$m^2 n$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.



What do we have to compute?

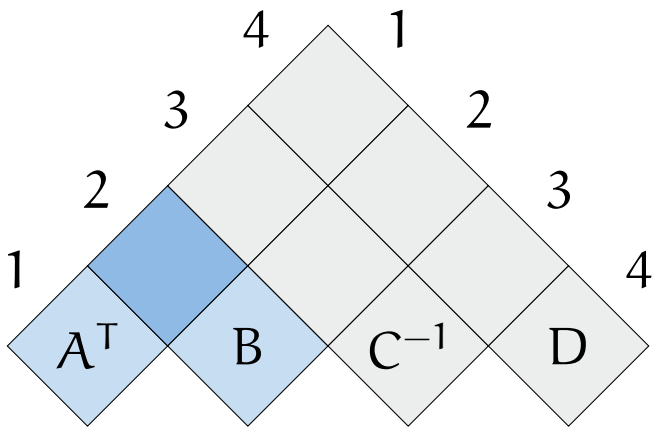
$$A^T \cdot B$$

Name	Pattern	Constraints	Cost [#FLOPs]
GEMM	$X^T Y$	-	$2mnk$
TRMM	$X^T Y$	is_lower_triangular(X)	$m^2 n$
TRMM	$X^T Y$	is_upper_triangular(X)	$m^2 n$
TRSM	$X^{-1} Y$	is_lower_triangular(X)	$m^2 n$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.



What do we have to compute?

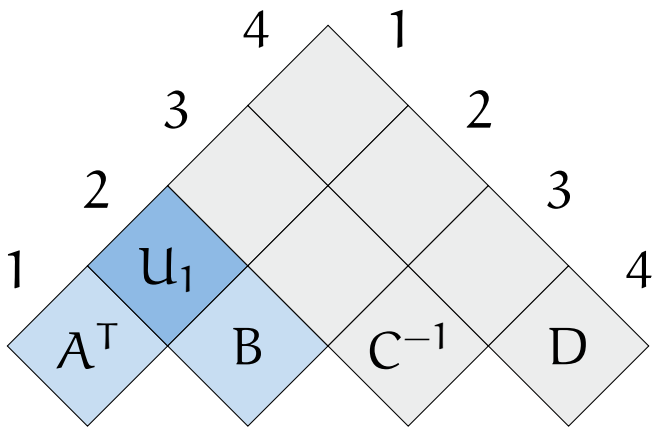
$$U_1 := A^T \cdot B$$

Name	Pattern	Constraints	Cost [#FLOPs]
GEMM	$X^T Y$	-	$2mnk$
TRMM	$X^T Y$	is_lower_triangular(X)	$m^2 n$
TRMM	$X^T Y$	is_upper_triangular(X)	$m^2 n$
TRSM	$X^{-1} Y$	is_lower_triangular(X)	$m^2 n$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.



What do we have to compute?

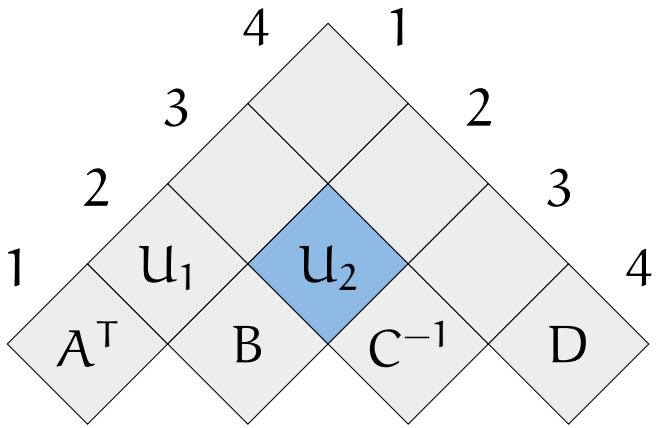
$$U_1 := A^T \cdot B$$

Name	Pattern	Constraints	Cost [#FLOPs]
GEMM	$X^T Y$	-	$2mnk$
TRMM	$X^T Y$	is_lower_triangular(X)	$m^2 n$
TRMM	$X^T Y$	is_upper_triangular(X)	$m^2 n$
TRSM	$X^{-1} Y$	is_lower_triangular(X)	$m^2 n$

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

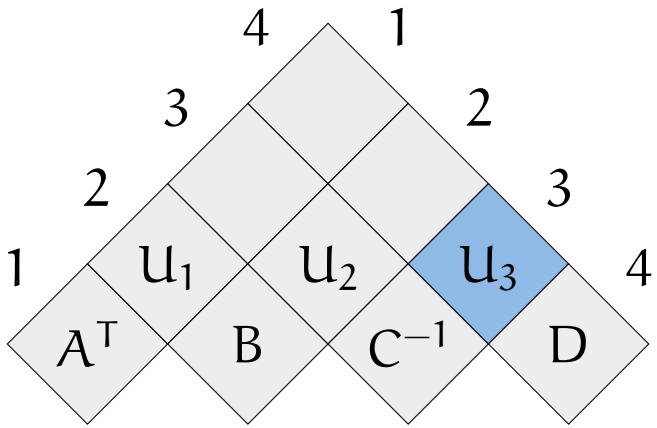
Example: $A^T B C^{-1} D$ A is lower triangular.



Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

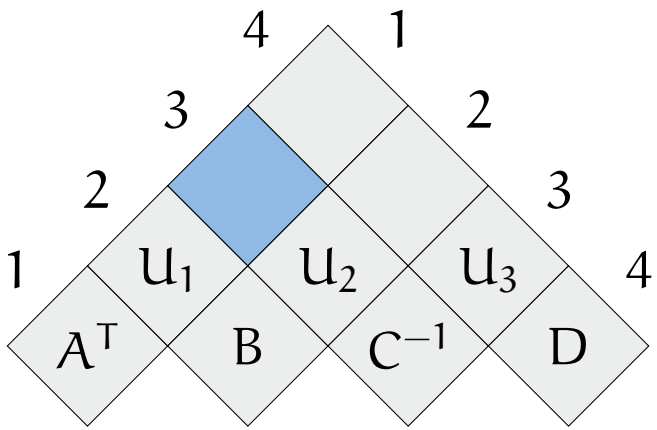
Example: $A^T B C^{-1} D$ A is lower triangular.



Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.

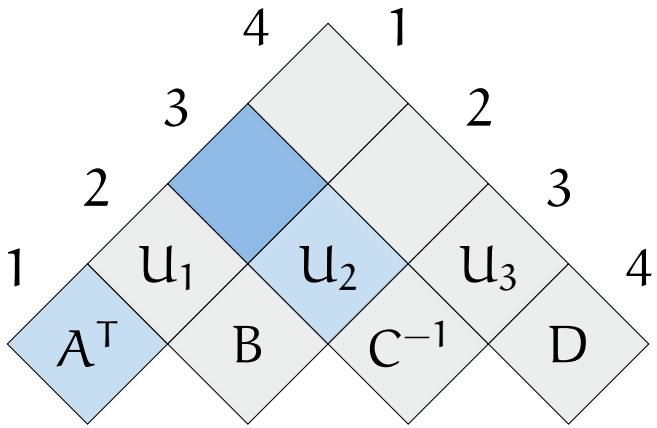


Expression	Kernel	Cost

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.

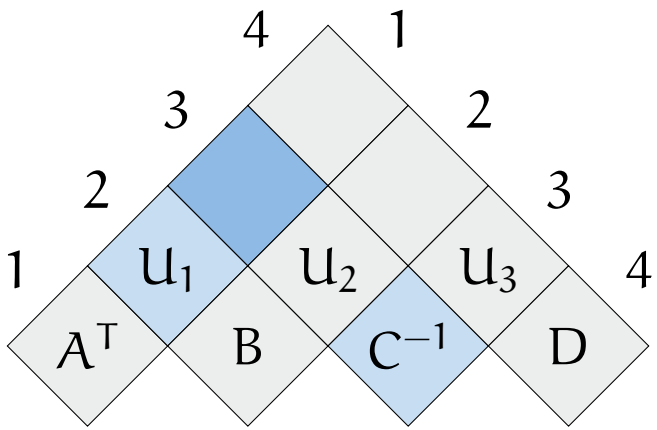


Expression	Kernel	Cost
$A^T \cdot U_2$	TRMM	20

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

Example: $A^T B C^{-1} D$ A is lower triangular.

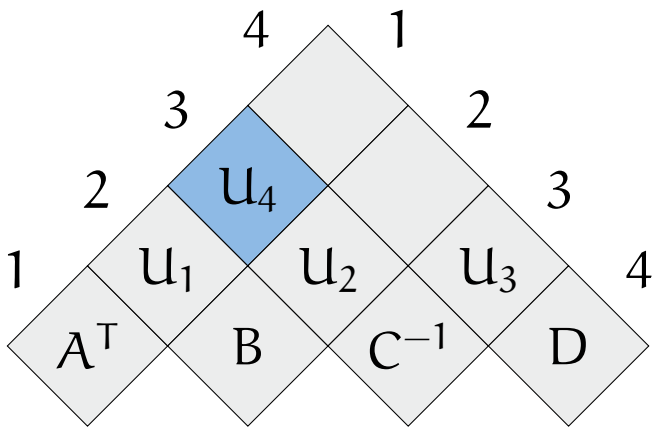


Expression	Kernel	Cost
$A^T \cdot U_2$	TRMM	20
$U_1 \cdot C^{-1}$	GESV	30

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

The Generalized Matrix Chain Algorithm

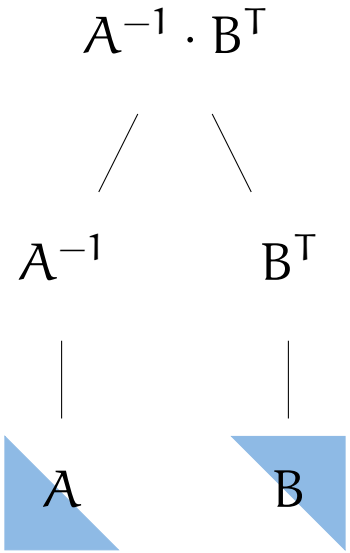
Example: $A^T B C^{-1} D$ A is lower triangular.



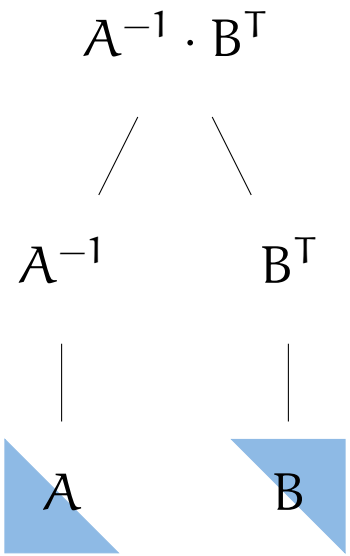
Expression	Kernel	Cost
$A^T \cdot u_2$	TRMM	20
$u_1 \cdot C^{-1}$	GESV	30

Code adapted from `https://tex.stackexchange.com/questions/110820/creating-a-triangular-table`

Inference of Properties

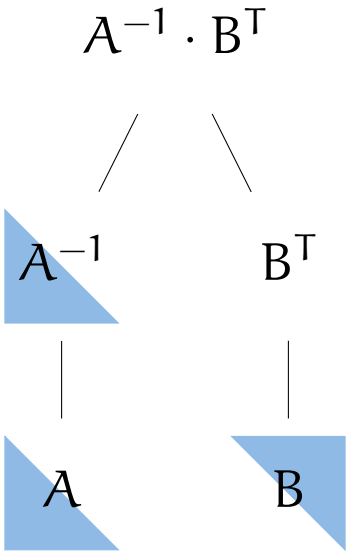


Inference of Properties



$\text{lowerTriangular}(X) \rightarrow \text{lowerTriangular}(X^{-1})$

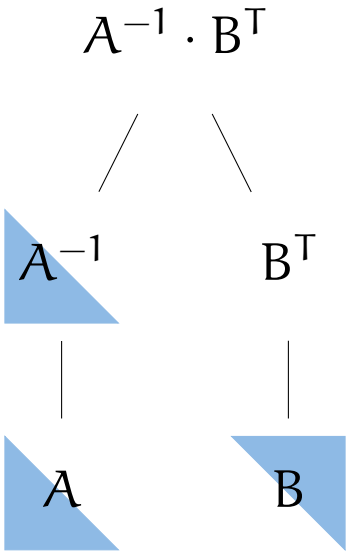
Inference of Properties



$\text{lowerTriangular}(X) \rightarrow \text{lowerTriangular}(X^{-1})$

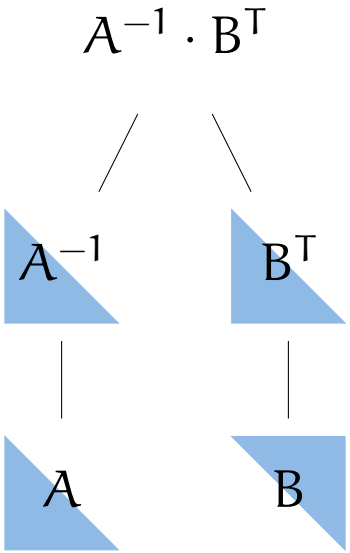
The Generalized Matrix Chain Algorithm

Inference of Properties



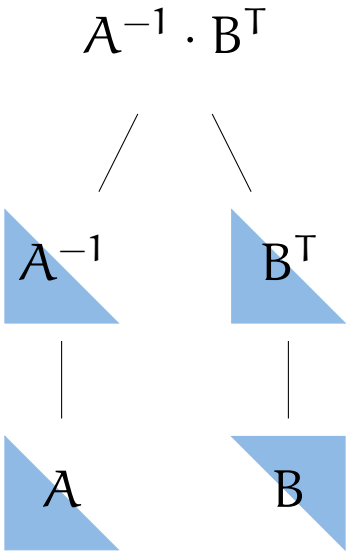
upperTriangular $(X) \rightarrow$ lowerTriangular (X^T)

Inference of Properties



upperTriangular $(X) \rightarrow$ lowerTriangular (X^T)

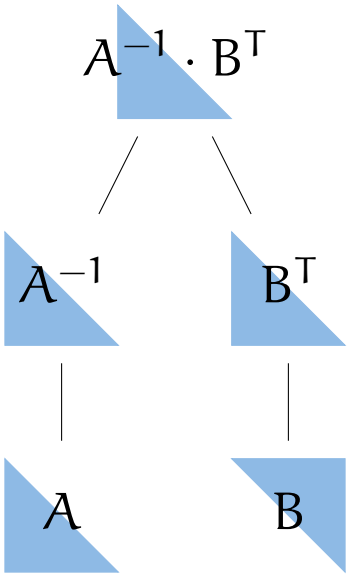
Inference of Properties



$$\text{lowerTriangular}(X) \wedge \text{lowerTriangular}(Y) \rightarrow \text{lowerTriangular}(X \cdot Y)$$

The Generalized Matrix Chain Algorithm

Inference of Properties



$$\text{lowerTriangular}(X) \wedge \text{lowerTriangular}(Y) \rightarrow \text{lowerTriangular}(X \cdot Y)$$

Experimental Results

- 100 randomly generated chains.
- 3 to 10 operands.
- We generate Julia code using BLAS/LAPACK wrappers
- We compare against
  - Julia
  - Matlab
  - Armadillo
  - Eigen
  - Blaze

Experimental Results

Naive

```
W = inv(A)*B*C*inv(D')*E*F
```

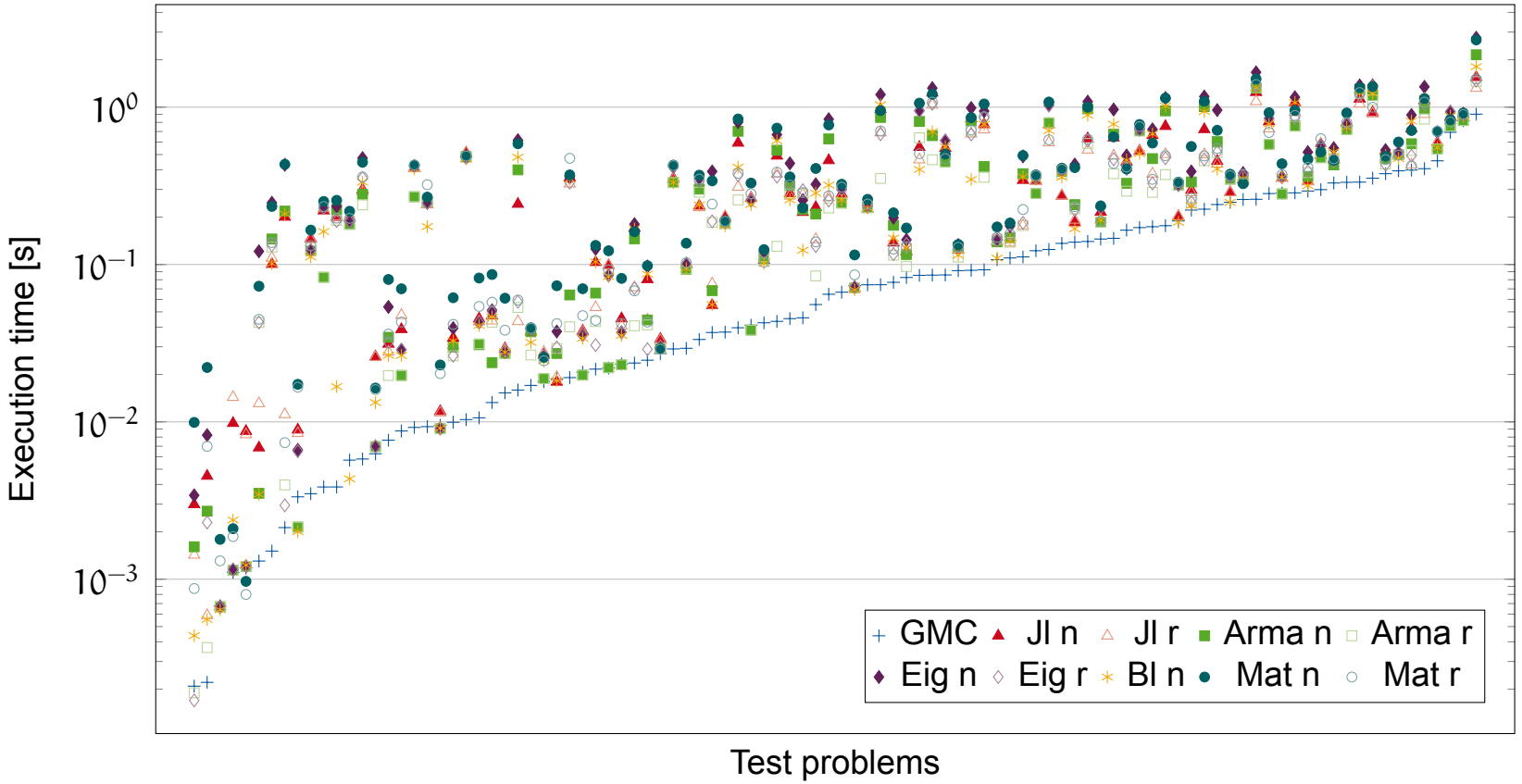
Recommended

```
W = (A\B)*C*(D'\E)*F
```

Generated

```
trsm!('L', 'U', 'T', 'N', 1.0, ml3, ml4)
ml6 = Array{Float64}(1200, 800)
gemm!('N', 'N', 1.0, ml2, ml4, 0.0, ml6)
ml7 = Array{Float64}(1500, 800)
gemm!('N', 'N', 1.0, ml1, ml6, 0.0, ml7)
posv!('L', ml0, ml7)
ml8 = Array{Float64}(1500, 1000)
gemm!('N', 'N', 1.0, ml7, ml5, 0.0, ml8)
```

Experimental Results



References

- Edward Anderson, Zhaojun Bai, et al.
  LAPACK Users' guide, volume 9.
  SIAM, 1999.
- Thomas H. Cormen, Ronald L. Rivest, and Charles E. Leiserson.
  Introduction to Algorithms.
  McGraw-Hill, Inc., 1990.
- Jack J. Dongarra, Jeremy Du Croz, et al.
  A set of Level 3 Basic Linear Algebra Subprograms.
  ACM TOMS, 16(1):1–17, 1990.