Knowledge-Based Linear Algebra Compiler

Henrik Barthels

High Performance and Automatic Computing Group
Advisor: Paolo Bientinesi

October 15, 2015

Introduction

$$b := (A^T A)^{-1} A^T y$$

Introduction

$$b := (A^T A)^{-1} A^T y \qquad\qquad T_1 := A^T A$$

Introduction

$$b := (A^T A)^{-1} A^T y = T_1^{-1} A^T y \qquad T_1 := A^T A$$

Introduction

$$
\begin{aligned}
b &:= (A^T A)^{-1} A^T y \\
&= T_1^{-1} A^T y
\end{aligned}
\qquad\qquad
\begin{aligned}
T_1 &:= A^T A \\
T_2 &:= T_1^{-1}
\end{aligned}
$$

Introduction

$$\begin{aligned} b &:= (A^T A)^{-1} A^T y \\ &= T_1^{-1} A^T y \\ &= T_2 A^T y \end{aligned}$$

$$\begin{aligned} T_1 &:= A^T A \\ T_2 &:= T_1^{-1} \end{aligned}$$

Introduction

$$
\begin{aligned}
b &:= (A^\top A)^{-1} A^\top y \\
&= T_1^{-1} A^\top y \\
&= T_2 A^\top y
\end{aligned}
\qquad\qquad
\begin{aligned}
T_1 &:= A^\top A \\
T_2 &:= T_1^{-1} \\
T_3 &:= T_2 A^\top
\end{aligned}
$$

Introduction

$$\begin{aligned} b &:= (A^T A)^{-1} A^T y \\ &= T_1^{-1} A^T y \\ &= T_2 A^T y \\ &= T_3 y \end{aligned}$$

$$\begin{aligned} T_1 &:= A^T A \\ T_2 &:= T_1^{-1} \\ T_3 &:= T_2 A^T \end{aligned}$$

Introduction

$$\begin{aligned} b &:= (A^T A)^{-1} A^T y \\ &= T_1^{-1} A^T y \\ &= T_2 A^T y \\ &= T_3 y \end{aligned}$$

$$\begin{aligned} T_1 &:= A^T A \\ T_2 &:= T_1^{-1} \\ T_3 &:= T_2 A^T \\ b &:= T_3 y \end{aligned}$$

Introduction

$$b := (A^\top A)^{-1} A^\top y$$

Introduction

$$b := (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}y \qquad\qquad QR := A$$

Introduction

$$b := (A^T A)^{-1} A^T y \qquad\qquad QR := A$$
$$= ((QR)^T QR)^{-1} (QR)^T y$$

Introduction

$$\begin{aligned} b &:= (A^T A)^{-1} A^T y \\ &= ((QR)^T QR)^{-1} (QR)^T y \\ &= (R^T Q^T QR)^{-1} R^T Q^T y \end{aligned} \qquad QR := A$$

Introduction

$$
\begin{aligned}
b &:= (A^T A)^{-1} A^T y \\
&= ((QR)^T QR)^{-1} (QR)^T y \\
&= (R^T Q^T QR)^{-1} R^T Q^T y \\
&= (R^T R)^{-1} R^T Q^T y
\end{aligned}
\qquad\qquad QR := A
$$

Introduction

$$\begin{aligned} b &:= (A^\top A)^{-1} A^\top y \\ &= ((QR)^\top QR)^{-1} (QR)^\top y \\ &= (R^\top Q^\top QR)^{-1} R^\top Q^\top y \\ &= (R^\top R)^{-1} R^\top Q^\top y \\ &= R^{-1} R^{-\top} R^\top Q^\top y \end{aligned}$$

$$QR := A$$

Introduction

$$\begin{aligned} b &:= (A^T A)^{-1} A^T y \\ &= ((QR)^T QR)^{-1} (QR)^T y \\ &= (R^T Q^T QR)^{-1} R^T Q^T y \\ &= (R^T R)^{-1} R^T Q^T y \\ &= R^{-1} R^{-T} R^T Q^T y \\ &= R^{-1} Q^T y \end{aligned}$$

$$QR := A$$

Introduction

$$
\begin{aligned}
b &:= (A^T A)^{-1} A^T y & QR &:= A \\
&= ((QR)^T QR)^{-1} (QR)^T y \\
&= (R^T Q^T QR)^{-1} R^T Q^T y \\
&= (R^T R)^{-1} R^T Q^T y \\
&= R^{-1} R^{-T} R^T Q^T y \\
&= R^{-1} Q^T y & t_1 &:= Q^T y
\end{aligned}
$$

Introduction

$$
\begin{aligned}
b &:= (A^T A)^{-1} A^T y & QR &:= A \\
&= ((QR)^T QR)^{-1} (QR)^T y \\
&= (R^T Q^T QR)^{-1} R^T Q^T y \\
&= (R^T R)^{-1} R^T Q^T y \\
&= R^{-1} R^{-T} R^T Q^T y \\
&= R^{-1} Q^T y & t_1 &:= Q^T y \\
&= R^{-1} t_1
\end{aligned}
$$

Introduction

$$
\begin{aligned}
b &:= (A^T A)^{-1} A^T y & QR &:= A \\
&= ((QR)^T QR)^{-1} (QR)^T y \\
&= (R^T Q^T QR)^{-1} R^T Q^T y \\
&= (R^T R)^{-1} R^T Q^T y \\
&= R^{-1} R^{-T} R^T Q^T y \\
&= R^{-1} Q^T y & t_1 &:= Q^T y \\
&= R^{-1} t_1 & b &:= R^{-1} t_1
\end{aligned}
$$

Using Linear Algebra Knowledge

Input

```
Matrix A <Symmetric, NonSingular>
Vector b
Vector x

x = inv(A) * b
```

Using Linear Algebra Knowledge

Input

```
Matrix A <Symmetric, NonSingular>
Vector b
Vector x

x = inv(A) * b
```

Knowledgebase

$$(A \cdot B \cdot \ldots)^{T} \rightarrow \ldots \cdot B^{T} \cdot A^{T}$$

$$A^{-1} \wedge \text{LowerTriangular}(A) \rightarrow \text{LowerTriangular}(A^{-1})$$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n \times n}, x, c \in \mathbb{R}^n$$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n \times n}, x, c \in \mathbb{R}^n$$

$x := (AB)c$ $\qquad$ $x := A(Bc)$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n \times n}, x, c \in \mathbb{R}^n$$

$$x := (AB)c \qquad\qquad x := A(Bc)$$

$$T := AB \qquad 2n^3$$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n \times n}, x, c \in \mathbb{R}^n$$

$$x := (AB)c \qquad\qquad x := A(Bc)$$

$$T := AB \qquad 2n^3$$

$$x := Tc \qquad 2n^2$$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n\times n}, x, c \in \mathbb{R}^n$$

$$x := (AB)c \qquad\qquad x := A(Bc)$$

$$T := AB \qquad 2n^3$$

$$x := Tc \qquad 2n^2$$

$$2n^3 + 2n^2$$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n\times n}, x, c \in \mathbb{R}^n$$

$$x := (AB)c$$

$$T := AB \qquad 2n^3$$

$$x := Tc \qquad 2n^2$$

$$2n^3 + 2n^2$$

$$x := A(Bc)$$

$$t := Bc \qquad 2n^2$$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n \times n}, x, c \in \mathbb{R}^n$$

$$x := (AB)c$$

$$T := AB \qquad 2n^3$$

$$x := Tc \qquad 2n^2$$

$$2n^3 + 2n^2$$

$$x := A(Bc)$$

$$t := Bc \qquad 2n^2$$

$$x := At \qquad 2n^2$$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n \times n}, x, c \in \mathbb{R}^n$$

$$x := (AB)c$$

$$T := AB \qquad 2n^3$$

$$x := Tc \qquad 2n^2$$

$$2n^3 + 2n^2$$

$$x := A(Bc)$$

$$t := Bc \qquad 2n^2$$

$$x := At \qquad 2n^2$$

$$4n^2$$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n \times n}, x, c \in \mathbb{R}^{n}, n = 200$$

$$x := (AB)c$$

$$T := AB \qquad 2n^3$$

$$x := Tc \qquad 2n^2$$

$$2n^3 + 2n^2$$

$$x := A(Bc)$$

$$t := Bc \qquad 2n^2$$

$$x := At \qquad 2n^2$$

$$4n^2$$

Matrix Chain Problem

$$x := ABc \qquad A, B \in \mathbb{R}^{n \times n}, x, c \in \mathbb{R}^n, n = 200$$

$x := (AB)c$		$x := A(Bc)$	
$T := AB$	$2n^3$	$t := Bc$	$2n^2$
$x := Tc$	$2n^2$	$x := At$	$2n^2$
$2n^3 + 2n^2$		$4n^2$	
16.080.000		160.000	

Matrix Chain Problem (indexed)

for i in $\{1, \ldots, m\}$:
    for j in $\{1, \ldots, k\}$:
        $x_{ij} := A_i B c_j$

$A_i, B \in \mathbb{R}^{n \times n}, x, c_j \in \mathbb{R}^n$

Matrix Chain Problem (indexed)

for i in $\{1, \dots, m\}$:
    for j in $\{1, \dots, k\}$:
        $x_{ij} := A_i B c_j$

$A_i, B \in \mathbb{R}^{n \times n}, x, c_j \in \mathbb{R}^n$

for i in $\{1, \dots, m\}$:
    for j in $\{1, \dots, k\}$:
        $T_i := A_i B$
        $x_{ij} := T_i c_j$

for i in $\{1, \dots, m\}$:
    for j in $\{1, \dots, k\}$:
        $t_j := B c_j$
        $x_{ij} := A_i t_j$

Matrix Chain Problem (indexed)

for i in $\{1, \ldots, m\}$:
    for j in $\{1, \ldots, k\}$:
        $x_{ij} := A_i B c_j$

$A_i, B \in \mathbb{R}^{n \times n}, x, c_j \in \mathbb{R}^n$

for i in $\{1, \ldots, m\}$:
    $T_i := A_i B$
    for j in $\{1, \ldots, k\}$:
        $x_{ij} := T_i c_j$

for i in $\{1, \ldots, m\}$:
    for j in $\{1, \ldots, k\}$:
        $t_j := B c_j$
        $x_{ij} := A_i t_j$

Matrix Chain Problem (indexed)

```
for i in {1, ..., m}:
    for j in {1, ..., k}:
        x_ij := A_i B c_j
```

$A_i, B \in \mathbb{R}^{n \times n}, x, c_j \in \mathbb{R}^n$

```
for i in {1, ..., m}:
    T_i := A_i B
    for j in {1, ..., k}:
        x_ij := T_i c_j
```

$$2mn^3 + 2mkn^2$$

```
for i in {1, ..., m}:
    for j in {1, ..., k}:
        t_j := B c_j
        x_ij := A_i t_j
```

Matrix Chain Problem (indexed)

```
for i in {1, ..., m}:
    for j in {1, ..., k}:
        x_ij := A_i B c_j
```

$A_i, B \in \mathbb{R}^{n \times n}, x, c_j \in \mathbb{R}^n$

```
for i in {1, ..., m}:
    T_i := A_i B
    for j in {1, ..., k}:
        x_ij := T_i c_j
```

$$2mn^3 + 2mkn^2$$

```
for j in {1, ..., k}:
    for i in {1, ..., m}:
        t_j := B c_j
        x_ij := A_i t_j
```

Matrix Chain Problem (indexed)

```
for i in {1, ..., m}:
    for j in {1, ..., k}:
        x_ij := A_i B c_j
```

$A_i, B \in \mathbb{R}^{n \times n}, x, c_j \in \mathbb{R}^n$

for i in $\{1, \ldots, m\}$:
    $T_i := A_i B$
    for j in $\{1, \ldots, k\}$:
        $x_{ij} := T_i c_j$

$$2mn^3 + 2mkn^2$$

for j in $\{1, \ldots, k\}$:
    $t_j := B c_j$
    for i in $\{1, \ldots, m\}$:
        $x_{ij} := A_i t_j$

Matrix Chain Problem (indexed)

```
for i in {1, ..., m}:
    for j in {1, ..., k}:
        x_ij := A_i B c_j
```

$A_i, B \in \mathbb{R}^{n \times n}, x, c_j \in \mathbb{R}^n$

```
for i in {1, ..., m}:
    T_i := A_i B
    for j in {1, ..., k}:
        x_ij := T_i c_j
```

$$2mn^3 + 2mkn^2$$

```
for j in {1, ..., k}:
    t_j := B c_j
    for i in {1, ..., m}:
        x_ij := A_i t_j
```

$$2kn^2 + 2mkn^2$$

Matrix Chain Problem (indexed)

for i in $\{1, \ldots, m\}$:
    for j in $\{1, \ldots, k\}$:
        $x_{ij} := A_i B c_j$

$A_i, B \in \mathbb{R}^{n \times n}, x, c_j \in \mathbb{R}^n$

for i in $\{1, \ldots, m\}$:
    $T_i := A_i B$
    for j in $\{1, \ldots, k\}$:
        $x_{ij} := T_i c_j$

$$2mn^3 + 2mkn^2$$

for j in $\{1, \ldots, k\}$:
    $t_j := B c_j$
    for i in $\{1, \ldots, m\}$:
        $x_{ij} := A_i t_j$

$$2kn^2 + 2mkn^2$$

Remark: Alternative Operations

$$X := AL^{-1}B$$

Common Subexpressions

$$X := ABCAB$$

Common Subexpressions

$$X := ABCAB \quad \rightarrow \quad \begin{aligned} U &:= AB \\ X &:= UCU \end{aligned}$$

Common Subexpressions

$$X := ABCAB \quad \rightarrow \quad \begin{aligned} U &:= AB \\ X &:= UCU \end{aligned}$$

$$X := A^T BCB^T A$$

Common Subexpressions

$$X := ABCAB \quad \rightarrow \quad \begin{aligned} U &:= AB \\ X &:= UCU \end{aligned}$$

$$X := A^T BCB^T A \quad \rightarrow \quad \begin{aligned} U &:= A^T B \\ X &:= UCU^T \end{aligned}$$

Common Subexpressions

$$X := ab^T ab^T \qquad a, b \in \mathbb{R}^n, X \in \mathbb{R}^{n \times n}$$

Common Subexpressions

$$X := ab^T ab^T \qquad a, b \in \mathbb{R}^n, X \in \mathbb{R}^{n \times n}$$

$$U := ab^T$$
$$X := UU$$

Common Subexpressions

$$X := ab^T ab^T \qquad a, b \in \mathbb{R}^n, X \in \mathbb{R}^{n \times n}$$

$$U := ab^T$$
$$X := UU$$

$$2n^2 + 2n^3$$

Common Subexpressions

$$X := ab^T ab^T \qquad a, b \in \mathbb{R}^n, X \in \mathbb{R}^{n \times n}$$

$$U := ab^T$$
$$X := UU$$

$$\alpha := b^T a$$
$$X := \alpha a b^T$$

$$2n^2 + 2n^3$$

Common Subexpressions

$$X := ab^T ab^T \qquad a, b \in \mathbb{R}^n, X \in \mathbb{R}^{n \times n}$$

$U := ab^T$	$\alpha := b^T a$
$X := UU$	$X := \alpha a b^T$
$2n^2 + 2n^3$	$3n + 2n^2$

Blocked Operands

for i in $\{1, \ldots, n\}$:

$$X_i := AB_i$$

Blocked Operands

for i in $\{1, \ldots, n\}$:

$$X_i := A \left(\frac{B'}{B'_i} \right)$$

Blocked Operands

for i in $\{1, \ldots, n\}$:

$$X_i := \left(\begin{array}{c|c} A_L & A_R \end{array} \right) \left(\begin{array}{c} B' \\ \hline B'_i \end{array} \right)$$

Blocked Operands

for i in $\{1, \ldots, n\}$:

$$X_i := A_L B' + A_R B'_i$$

Blocked Operands

$T := A_L B'$
for i in $\{1, \ldots, n\}$:

$\quad X_i := T + A_R B'_i$

Inverse Operator

$$X := (S + \alpha I)^{-1} \qquad S \in \mathbb{R}^{n \times n}, \text{ symmetric}$$
$$\alpha \in \mathbb{R}$$

Inverse Operator

$$X := (S + \alpha I)^{-1} \qquad S \in \mathbb{R}^{n \times n}, \text{symmetric}$$
$$\alpha \in \mathbb{R}$$

$$(S + \alpha I)^{-1} = (ZWZ^{\mathsf{T}} + \alpha I)^{-1}$$

Inverse Operator

$$X := (S + \alpha I)^{-1} \qquad S \in \mathbb{R}^{n \times n}, \text{symmetric}$$

$$\alpha \in \mathbb{R}$$

$$(S + \alpha I)^{-1} = (ZWZ^{\mathsf{T}} + \alpha I)^{-1} \qquad ZZ^{\mathsf{T}} = I$$

Inverse Operator

$$X := (S + \alpha I)^{-1} \qquad S \in \mathbb{R}^{n \times n}, \text{ symmetric}$$
$$\alpha \in \mathbb{R}$$

$$\begin{aligned}(S + \alpha I)^{-1} &= (ZWZ^{\mathsf{T}} + \alpha I)^{-1} \qquad ZZ^{\mathsf{T}} = I \\ &= (ZWZ^{\mathsf{T}} + \alpha ZZ^{\mathsf{T}})^{-1}\end{aligned}$$

Inverse Operator

$$X := (S + \alpha I)^{-1} \qquad S \in \mathbb{R}^{n \times n}, \text{ symmetric}$$
$$\alpha \in \mathbb{R}$$

$$\begin{aligned} (S + \alpha I)^{-1} &= (ZWZ^{\mathsf{T}} + \alpha I)^{-1} \qquad\qquad ZZ^{\mathsf{T}} = I \\ &= (ZWZ^{\mathsf{T}} + \alpha ZZ^{\mathsf{T}})^{-1} \\ &= (Z(W + \alpha I)Z^{\mathsf{T}})^{-1} \end{aligned}$$

Inverse Operator

$$X := (S + \alpha I)^{-1} \qquad S \in \mathbb{R}^{n \times n}, \text{ symmetric}$$
$$\alpha \in \mathbb{R}$$

$$\begin{aligned}(S + \alpha I)^{-1} &= (ZWZ^T + \alpha I)^{-1} \qquad ZZ^T = I \\ &= (ZWZ^T + \alpha ZZ^T)^{-1} \\ &= (Z(W + \alpha I)Z^T)^{-1} \\ &= Z^{-T}(W + \alpha I)^{-1} Z^{-1}\end{aligned}$$

Inverse Operator

$$X := (S + \alpha I)^{-1} \qquad S \in \mathbb{R}^{n \times n}, \text{symmetric}$$
$$\alpha \in \mathbb{R}$$

$$\begin{aligned}
(S + \alpha I)^{-1} &= (ZWZ^T + \alpha I)^{-1} & ZZ^T &= I \\
&= (ZWZ^T + \alpha ZZ^T)^{-1} \\
&= (Z(W + \alpha I)Z^T)^{-1} \\
&= Z^{-T}(W + \alpha I)^{-1} Z^{-1} & Z^{-1} &= Z^T
\end{aligned}$$

Inverse Operator

$$X := (S + \alpha I)^{-1} \qquad S \in \mathbb{R}^{n \times n}, \text{ symmetric}$$
$$\alpha \in \mathbb{R}$$

$$\begin{aligned}
(S + \alpha I)^{-1} &= (ZWZ^T + \alpha I)^{-1} & ZZ^T = I \\
&= (ZWZ^T + \alpha ZZ^T)^{-1} \\
&= (Z(W + \alpha I)Z^T)^{-1} \\
&= Z^{-T}(W + \alpha I)^{-1} Z^{-1} & Z^{-1} = Z^T \\
&= Z(W + \alpha I)^{-1} Z^T
\end{aligned}$$

Conclusion

- Generating solutions is just one half of the problem.
- Productivity tool.
- Implementation
  - Python.
  - Output language?
  - Integration into existing software?
- Example problems.