Inverses

Stephen Boyd

EE103
Stanford University

October 27, 2015
Left and right inverses

Inverse

Solving linear equations

Examples

Pseudo-inverse
Left inverses

- a number $x$ that satisfies $xa = 1$ is called the inverse of $a$
- inverse (i.e., $1/a$) exists if and only if $a \neq 0$, and is unique
- a matrix $X$ that satisfies $XA = I$ is called a left inverse of $A$ (and we say that $A$ is left-invertible)
- example: the matrix

$$A = \begin{bmatrix} -3 & -4 \\ 4 & 6 \\ 1 & 1 \end{bmatrix}$$

has two different left inverses:

$$B = \frac{1}{9} \begin{bmatrix} -11 & -10 & 16 \\ 7 & 8 & -11 \end{bmatrix}, \quad C = \frac{1}{2} \begin{bmatrix} 0 & -1 & 6 \\ 0 & 1 & -4 \end{bmatrix}$$
if $A$ has a left inverse $C$ then the columns of $A$ are independent

To see this: if $Ax = 0$ and $CA = I$ then

$$0 = C0 = C(Ax) = (CA)x = Ix = x$$

we’ll see later the converse is also true, so

*a matrix is left-invertible if and only if its columns are independent*

so left-invertible matrices are tall or square
Solving linear equations with a left inverse

- suppose $Ax = b$, and $A$ has a left inverse $C$
- then $Cb = C(Ax) = (CA)x = Ix = x$
- so multiplying the right-hand side by a left inverse yields the solution
Right inverses

- A matrix $X$ that satisfies $AX = I$ is a right inverse of $A$ (and we say that $A$ is right-invertible)
- $A$ is right-invertible if and only if $A^T$ is left-invertible:
  \[
  AX = I \iff (AX)^T = I \iff X^T A^T = I
  \]
- so we conclude
  \textit{A is right-invertible if and only if its rows are linearly independent}
- right-invertible matrices are wide or square
Solving linear equations with a right inverse

▶ suppose $A$ has a right inverse $B$
▶ consider the (square or underdetermined) equations $Ax = b$
▶ $x = Bb$ is a solution:

$$Ax = A(Bb) = (AB)b = Ib = b$$

▶ so $Ax = b$ has a solution for any $b$
Outline

Left and right inverses

Inverse

Solving linear equations

Examples

Pseudo-inverse
if $A$ has a left and a right inverse, they are unique and equal (and we say that $A$ is invertible)

so $A$ must be square

to see this: if $AX = I$, $YA = I$

$$X = (YA)X = Y(AX) = Y$$

we denote them by $A^{-1}$: $A^{-1}A = AA^{-1} = I$

inverse of inverse: $(A^{-1})^{-1} = A$
Solving square systems of linear equations

- suppose $A$ is invertible
- for any $b$, $Ax = b$ has the unique solution $x = A^{-1}b$
- matrix generalization of simple scalar equation $ax = b$ having solution $x = (1/a)b$ (for $a \neq 0$)
Invertible matrices

the following are equivalent for a square matrix $A$:

- $A$ is invertible
- columns of $A$ are linearly independent
- rows of $A$ are linearly independent
- $A$ has a left inverse
- $A$ has a right inverse

if any of these hold, all others do
Examples

- $I^{-1} = I$
- if $Q$ is orthogonal, \textit{i.e.}, $Q^TQ = I$, then $Q^{-1} = Q^T$
- $2 \times 2$ matrix $A$ is invertible if and only if $A_{11}A_{22} \neq A_{12}A_{21}$

$$A^{-1} = \frac{1}{A_{11}A_{22} - A_{12}A_{21}} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}$$

- you need to know this formula
- there are similar but \textit{much} more complicated formulas for larger matrices (and no, you do not need to know them)
Properties

\[ (AB)^{-1} = B^{-1}A^{-1} \] (provided inverses exist)

\[ (A^T)^{-1} = (A^{-1})^T \] (sometimes denoted \( A^{-T} \))

Negative matrix powers: \( (A^{-1})^k \) is denoted \( A^{-k} \)

With \( A^0 = I \), identity \( A^k A^l = A^{k+l} \) holds for any integers \( k, l \)
Triangular matrices

- lower triangular $L$ with nonzero diagonal entries is invertible
- so see this, write $Lx = 0$ as

\[
\begin{align*}
L_{11}x_1 &= 0 \\
L_{21}x_1 + L_{22}x_2 &= 0 \\
&\vdots \\
L_{n1}x_1 + L_{n2}x_2 + \cdots + L_{n,n-1}x_{n-1} + L_{nn}x_n &= 0
\end{align*}
\]

- from first equation, $x_1 = 0$ (since $L_{11} \neq 0$)
- second equation reduces to $L_{22}x_2 = 0$, so $x_2 = 0$ (since $L_{22} \neq 0$)
- and so on

- upper triangular $R$ with nonzero diagonal entries is invertible
Outline

Left and right inverses

Inverse

Solving linear equations

Examples

Pseudo-inverse
suppose $R$ is upper triangular with nonzero diagonal entries
write out $Rx = b$ as
\[
\begin{align*}
R_{11}x_1 + R_{12}x_2 + \cdots + R_{1,n-1}x_{n-1} + R_{1n}x_n &= b_1 \\
& \vdots \\
R_{n-1,n-1}x_{n-1} + R_{n-1,n}x_n &= b_{n-1} \\
R_{nn}x_n &= b_n
\end{align*}
\]
from last equation we get $x_n = b_n/R_{nn}$
from 2nd to last equation we get
\[
x_{n-1} = \frac{(b_{n-1} - R_{n-1,n}x_n)}{R_{n-1,n-1}}
\]
continue to get $x_{n-2}, x_{n-3}, \ldots, x_1$
Back substitution

- called *back substitution* since we find the variables in reverse order, substituting the already known values of \( x_i \)
- computes \( x = R^{-1}b \)
- complexity:
  - first step requires 1 flop (division)
  - 2nd step needs 3 flops
  - \( i \)th step needs \( 2i - 1 \) flops

Total is \( 1 + 3 + \cdots + (2n - 1) = n^2 \) flops
Solving linear equations via QR factorization

- assuming $A$ is invertible, let’s solve $Ax = b$, i.e., compute $x = A^{-1}b$
- columns of $A$ are linearly independent, so its QR factorization exists:
  
  - $A = QR$
  - $Q$ is orthogonal, so $Q^{-1} = Q^T$
  - $R$ is upper triangular with positive diagonal entries, hence invertible

- so $A^{-1} = (QR)^{-1} = R^{-1}Q^T$
- compute $x = R^{-1}(Q^Tb)$ by back substitution
Solving linear equations via QR factorization

given an $n \times n$ invertible matrix $A$ and an $n$-vector $b$

1. **QR factorization.** Compute the QR factorization $A = QR$.
2. Compute $Q^T b$.
3. **Back substitution.** Solve the triangular equation $Rx = Q^T b$ using back substitution.

- complexity $2n^3$ (step 1), $2n^2$ (step 2), $n^2$ (step 3)
- total is $2n^3 + 3n^2 \approx 2n^3$
Multiple right-hand sides

- let’s solve \( Ax_i = b_i, \ i = 1, \ldots, k, \) with \( A \) invertible
- carry out QR factorization once (\( 2n^3 \) flops)
- for \( i = 1, \ldots, k, \) solve \( Rx_i = Q^T b_i \) via back substitution (\( 3kn^2 \) flops)
- total is \( 2n^3 + 2kn^2 \) flops
- if \( k \) is small compared to \( n, \) same cost as solving one set of equations
Outline

Left and right inverses

Inverse

Solving linear equations

Examples

Pseudo-inverse
Polynomial interpolation

- let's find coefficients of a cubic polynomial

\[ p(x) = c_1 + c_2 x + c_3 x^2 + c_4 x^3 \]

that satisfies

\[ p(-1.1) = b_1, \quad p(-0.4) = b_2, \quad p(0.1) = b_3, \quad p(0.8) = b_4 \]

- write as \( Ac = b \), with

\[
A = \begin{bmatrix}
1 & -1.1 & (-1.1)^2 & (-1.1)^3 \\
1 & -0.4 & (-0.4)^2 & (-0.4)^3 \\
1 & 0.1 & (0.1)^2 & (0.1)^3 \\
1 & 0.8 & (0.8)^2 & (0.8)^3
\end{bmatrix}
\]
Polynomial interpolation

- (unique) coefficients given by \( c = A^{-1}b \), with

\[
A^{-1} = \begin{bmatrix}
-0.0201 & 0.2095 & 0.8381 & -0.0276 \\
0.1754 & -2.1667 & 1.8095 & 0.1817 \\
0.3133 & 0.4762 & -1.6667 & 0.8772 \\
-0.6266 & 2.381 & -2.381 & 0.6266
\end{bmatrix}
\]

- so, e.g., \( c_1 \) is not very sensitive to \( b_1 \) or \( b_4 \)
Outline

Left and right inverses

Inverse

Solving linear equations

Examples

Pseudo-inverse
Invertibility of Gram matrix

- $A$ has independent columns $\iff A^T A$ is invertible
- to see this, we'll show that $Ax = 0 \iff A^T Ax = 0$
- $\Rightarrow$: if $Ax = 0$ then $(A^T A)x = A^T (Ax) = A^T 0 = 0$
- $\Leftarrow$: if $(A^T A)x = 0$ then
  \[ 0 = x^T (A^T A)x = (Ax)^T (Ax) = \|Ax\|^2 = 0 \]
  so $Ax = 0$
Pseudo-inverse of tall matrix

- the \textit{pseudo-inverse} of $A$ with independent columns is

$$A^\dagger = (A^T A)^{-1} A^T$$

- it is a left inverse of $A$:

$$A^\dagger A = (A^T A)^{-1} A^T A = (A^T A)^{-1}(A^T A) = I$$

(we’ll soon see that it’s a very important left inverse of $A$)

- reduces to $A^{-1}$ when $A$ is square:

$$A^\dagger = (A^T A)^{-1} A^T = A^{-1} A^{-T} A^T = A^{-1} I = A^{-1}$$
Pseudo-inverse of wide matrix

- if $A$ is wide, with independent rows, $AA^T$ is invertible
- pseudo-inverse is defined as
  \[ A^\dagger = A^T (AA^T)^{-1} \]
- $A^\dagger$ is a right inverse of $A$:
  \[ AA^\dagger = AA^T (AA^T)^{-1} = I \]
  (we’ll see later it is an important right inverse)
- reduces to $A^{-1}$ when $A$ is square:
  \[ A^T (AA^T)^{-1} = A^T A^{-T} A^{-1} = A^{-1} \]
Pseudo-inverse via QR factorization

▶ suppose $A$ has independent columns, $A = QR$
▶ then $A^T A = (QR)^T (QR) = R^T Q^T Q R = R^T R$
▶ so

$$A^\dagger = (A^T A)^{-1} A^T = (R^T R)^{-1} (QR)^T = R^{-1} R^{-T} R^T Q^T = R^{-1} Q^T$$

▶ can compute $A^\dagger$ using back substitution on columns of $Q^T$

▶ for $A$ with independent rows, $A^\dagger = QR^{-T}$