Matrix facts

- Completion of squares
- Block LDU matrix decomposition
- Inverse of a block matrix
- Inverse of a sum
- Useful matrix identities
- Push-through identity

Completion of squares

- the *completion of squares* formula for quadratic polynomials is

$$ax^2 + 2bxy + dy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(d - \frac{b^2}{a}\right)y^2$$

- when $a > 0$, this tells us the *minimum with respect to x for fixed y*

$$\min_{x \in \mathbb{R}} \; ax^2 + 2bxy + dy^2 = \left(d - \frac{b^2}{a}\right)y^2$$

  which is achieved when $x = -\frac{b}{a}y$.

- this also gives a test for *global positivity*:

$$ax^2 + 2bxy + dy^2 > 0 \quad \text{for all nonzero } x, y \in \mathbb{R} \quad \iff \quad a > 0 \text{ and } d - \frac{b^2}{a} > 0$$

Completion of squares for matrices

- if $A \in \mathbb{R}^{n\times n}$ and $D \in \mathbb{R}^{m\times m}$ are symmetric matrices and $B \in \mathbb{R}^{n\times m}$, then

$$\begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \big(x + A^{-1}By\big)^T A \big(x + A^{-1}By\big) + y^T(D - B^T A^{-1} B)y$$

- compare with

$$ax^2 + 2bxy + dy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(d - \frac{b^2}{a}\right)y^2$$

- gives a general formula for quadratic optimization; if $A > 0$, then

$$\min_x \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = y^T(D - B^T A^{-1} B)y$$

  and the minimizing x is $\quad x_{\text{opt}} = -A^{-1}By$

Block LDU matrix decomposition

- the completion of squares formula gives a useful matrix decomposition

$$\begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} A & B \\ B^T & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \big(x + A^{-1}By\big)^T A \big(x + A^{-1}By\big)^T + y^T(D - B^T A^{-1} B)y$$

$$= \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} I & 0 \\ B^T A^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D - B^T A^{-1} B \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- since this holds for all x, y,

$$\begin{bmatrix} A & B \\ B^T & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ B^T A^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D - B^T A^{-1} B \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}$$

holds whenever A is invertible

Inverse of a block matrix

- the matrix $S = D - CA^{-1}B$ is called the *Schur complement* of A

- then $\begin{bmatrix} A & B \\ B^\mathsf{T} & D \end{bmatrix} > 0$ if and only if

$$A > 0 \qquad \text{and} \qquad D - B^\mathsf{T} A^{-1} B > 0$$

Inverse of a block matrix

- also holds for asymmetric matrices

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}$$

- this decomposition is easy to invert

$$\begin{aligned} \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} &= \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} A^{-1} & 0 \\ 0 & (D - CA^{-1}B)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix}^{-1} \\ &= \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & (D - CA^{-1}B)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \\ &= \begin{bmatrix} A^{-1} + A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}CA^{-1} & S^{-1} \end{bmatrix} \end{aligned}$$

Inverse of a sum

- we can also complete the square to minimize w.r.t. y instead of x, which gives another formula, which holds whenever D is invertible

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} T^{-1} & -T^{-1}BD^{-1} \\ -D^{-1}CT^{-1} & D^{-1} + D^{-1}CT^{-1}BD^{-1} \end{bmatrix}$$

where $T = A - BD^{-1}C$

- the two formulas for $\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1}$ must be equal, so

$$(A - BD^{-1}C)^{-1} = A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1}$$

called the *matrix inversion lemma* or *Sherman-Morrison-Woodbury* formula

Useful matrix identities

$$A(I+A)^{-1} = I - (I+A)^{-1}$$

because $(I+A)(I+A)^{-1} = I$ so

$$(I+A)^{-1} + A(I+A)^{-1} = I$$

More useful matrix identities

$$(I + AB)^{-1} = I - A(I + BA)^{-1}B$$

verify this directly; we have

$$\begin{aligned}\left(I - A(I + BA)^{-1}B\right)(I + AB) &= I + AB - A(I + BA)^{-1}B(I + AB) \\ &= I + AB - A(I + BA)^{-1}(I + BA)B \\ &= I\end{aligned}$$

- $I + AB$ is invertible if and only if $I + BA$ is
- true for any A and B, not just square ones

And more useful matrix identities

$$A(I + BA)^{-1} = (I + AB)^{-1}A$$

because $B(I + AB) = (I + BA)B$

- called *push-through* identity
- A on the left pushes in, and pushes out A on the right