Topics of the week

1. Pseudo-inverse, definition and properties;

2. Pseudo-inverse and minimum norm solution;

3. Pseudo-inverse and projection;

4. Polyhedron, projections of sets, Farkas lemma.

Left inverse of $A: U \to V$ is the matrix $A^\dagger$ s.t. $A^\dagger A = I$.
Interpretation: Given $y \in \operatorname{im} A$, find the **unique** pre-image $x = A^\dagger y$.
Criterion: There is $A^\dagger \iff A$ is injective $\iff$ columns of A are linearly independent.
Reason: If it's not injective, the pre-image of $y = Ax$ is non-unique.
Description: $(AA^\dagger)^2 = AA^\dagger$ is a projection on $\operatorname{im} A$.

Left pseudoinverse is $A^\dagger = (A^\top A)^{-1} A^\top$. For each y, finds x that minimizes $\|Ax - y\|$.
Extra: $AA^\dagger$ is the orthogonal projection on $\operatorname{im} A$.

Right inverse of $A: U \to V$ is the matrix $A^\dagger$ s.t. $AA^\dagger = I$.
Interpretation: Given $y \in V$, find **any** pre-image $x = A^\dagger y$.
Criterion: There is $A^\dagger \iff A$ is surjective $\iff$ rows of A are linearly independent.
Reason: If it's not surjective, there are y that without pre-image.
Description: $(A^\dagger A)^2 = A^\dagger A$ is a projection on a "representative" subspace $\operatorname{im} A^\dagger \subset U$.

Right pseudoinverse is $A^\dagger = A^\top (AA^\top)^{-1}$. For each y, finds x s.t. $Ax = y$ and $\|x\| \to \min$.
Extra: $A^\dagger A$ is the orthogonal projection on $(\ker A)^\perp = \operatorname{im} A^\top$.

Pseudoinverse of a matrix $A = CR$ is $A^\dagger = R^\dagger C^\dagger = R^\top (C^\top A R^\top)^{-1} C^\top$. Solves the problem:

$$\begin{gathered} \|x\| \to \min, \\ \text{s.t. } \|Ax - y\| \to \min. \end{gathered}$$

Implied equations Let $Ax = b$, then $(y^\top A)x = y^\top b$ for any y.

Fredholm theorem For a matrix A and a vector b, exactly one of the following holds:

1. There is x s.t. $Ax = b$;

2. There is y s.t. $y^\top A = 0$ and $y^\top b \neq 0$.

Interpretation: $Ax = b$ is infeasible $\iff$ there is an infeasible implied equation.
Proof: If there is such x, it means $y^\top b = y^\top (b - Ax) = 0$ for any y s.t. $y^\top A = 0$. If there is no such x, use Gaussian elimination to find y.
Variant: if there is no such x, use $y = Ax - b$, where $\|Ax - b\| \to \min$.

Conic combination of $v_1, \dots, v_n$ is $x_1 v_1 + \dots + x_n v_n$, where $x_1, \dots, x_n \geq 0$.

Implied inequalities Let $Ax \leq b$ and $y \geq 0$, then $y^\top Ax \leq y^\top b$.
Conic combinations: When $x \geq 0$, the linear combination $x_1 v_1 + \dots + x_n v_n$ is called **conic**.

Fourier-Motzkin elimination Consider a set of inequalities $Ax \leq b$. We can eliminate x_n:

$$\begin{cases} x_n \geq b_i - a_i^\top x', \\ x_n \leq b_j - a_j^\top x' \end{cases} \implies b_j - a_j^\top x' \leq x_n \leq b_i - a_i^\top x'$$

Add an inequality without x_n for each such (i, j). If there is a solution $x = (x', x_n)$, then it also satisfies the new system. Correspondingly, if there is a solution x' of the new system, it's possible to find x_n s.t. $x = (x', x_n)$ is the solution to the old system.

Note: Each new inequality is an implied inequality of the previous ones.

Note: Very inefficient in practice, compared to standard linear programming solvers.

Farkas lemma For a matrix A and a vector b, exactly one of the following holds:

1. There is x s.t. $Ax \leq b$;
2. There is $y \geq 0$ s.t. $y^\top A = 0$ and $y^\top b < 0$.

Interpretation: $Ax \leq b$ is infeasible $\iff$ there is an infeasible implied inequality.

Proof: If there is such x, it means $y^\top b = y^\top (b - Ax) \geq 0$ for any $y \geq 0$ s.t. $y^\top A = 0$. If there is no such x, use Fourier-Motzkin elimination to find y.

Variant: If there is no such x, use $y = Ax - b'$, where $\|Ax - b'\| \to \min$ s.t. $b' \leq b$.

In-class exercises Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ be arbitrary matrices.

1. Prove that if $\operatorname{rank} A = \operatorname{rank} B = n$, then $(AB)^\dagger = B^\dagger A^\dagger$.
2. Prove that $A^\dagger A A^\dagger = A^\dagger$.
3. Prove that $(A^\top)^\dagger = (A^\dagger)^\top$.
4. Prove that $A^\dagger A$ is symmetric and is the projection matrix for $\operatorname{im} A^\top$.