Lecture #2: Geometry and Separability

Recall:

def: The nonnegative rank

$$\text{rank}^+(M) = \text{smallest } r \text{ s.t. } \underset{m\times n}{M} = \underset{m\times r}{A}\,\underset{r\times n}{W}$$

where $A, W \geq 0$

Geometric picture

def: The cone C_A generated by colns of A is

$$C_A = \{Ax \mid x \geq 0\}$$

Claim: Given $M \geq 0$ and $A \geq 0$

$\exists W \geq 0$ s.t. $M = AW$

$$\Updownarrow$$

$$C_M \subseteq C_A$$

Proof: (By definitions)

$(\Rightarrow)$ If $M = AW$ then $M_i = AW_i \in C_A$,

similarly for all nonnegative linear combinations of M_i's

$(\Leftarrow)$ $C_M \subseteq C_A \Rightarrow \forall i \; M_i \in C_A$

$\Rightarrow \forall i \; M_i = A x_i$

Now set $W = [x_1, \ldots]$ $\boxtimes$

Given $M, A \geq 0$ it is easy to find W because it's a linear program

$$M = AW$$

$$W \geq 0$$

Can also write a convex program for approximating W

$$\min \|M - AW\|_F$$

$$\text{s.t. } W \geq 0$$

Leads to natural heuristic

Alternating Minimization

Guess $A \geq 0$

Repeat

$W \leftarrow \operatorname{argmin} \|M - AW\|_F \text{ s.t. } W \geq 0$

$A \leftarrow \operatorname{argmin} \|M - AW\|_F \text{ s.t. } A \geq 0$

Does it find a good solution?

A positive answer is predicated on identifying some tractable subclass

---

<u>Why</u> is NMF hard?

Hard instances are often <u>brittle</u>

① non unique solutions

② lack of robustness

Following [Vavasis], consider

<u>P1</u>: Given $M \geq 0$, is $\text{rank}^+(M) = \text{rank}(M)$?

<u>P2</u> Given $M \geq 0$ with $\text{rank}(M) = r$ and

$$M = \underbrace{\underset{m \times r}{U} \; \underset{r \times n}{V}}_{\text{arbitrary}}$$

Is there an invertible T s.t. $UT, T^{-1}V \geq 0$?

<u>Lemma</u>: P1 and P2 are equivalent

We want to show if answer to P1 is yes, then the answer to P2 (for any valid U, V) is yes too

Fact: If rank(M) = r and

$$M = UV \quad \text{and} \quad M = AW$$

are two factorizations with inner-dimension r then

① $\text{colspan}(U) = \text{colspan}(A) = \text{colspan}(M)$

② $\text{rowspan}(V) = \text{rowspan}(W) = \text{rowspan}(M)$

Proof: It suffices to prove ① b/c ② then follows by taking the transpose

By definition $\overset{(M=AW)}{}$ $\text{colspan}(M) \subseteq \text{colspan}(A)$, and since they both have dimension r they must be equal ☒

Proof of Lemma:

$$\text{colspan}(U) = \text{colspan}(A)$$

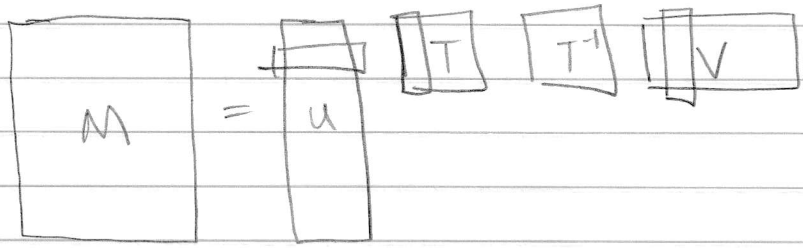
$$\Downarrow$$

$\exists$ invertible T s.t. $UT = A$

Then $M = \underbrace{UT}_{A} \; \underbrace{T^{-1}V}_{X}$

But this linear system has a unique soln in X b/c A has a left inverse so we must have $X = W$ ☒

Now let's interpret P2



Let $u_1, \ldots u_m$ be rows of U,
$t_1, \ldots t_r$ be colns of T
$v_1 \ldots v_n$ be colns of V

Also let $Q = \{x \mid u_i^T x \geq 0 \ \forall i\}$, also a cone

Fact: $UT \geq 0$ iff $C_T \subseteq Q$

Why? Again, proof follows by definition

Fact: $T^{-1} V \geq 0$ iff $C_V \subseteq C_T$

Why? $T^{-1} v_i$ are the coordinates of representing v_i in the basis of the t_j's

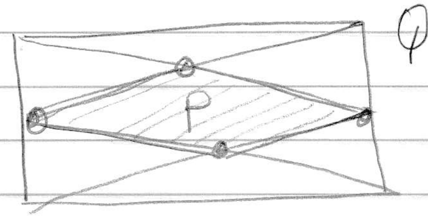
Hence P2 is equivalent to intermediate cone problem:

Given cones $P \subseteq Q$ is there a cone T s.t.

① T is generated by r vectors

② $P \subseteq T \subseteq Q$

Now consider the following gadget (cone ⇒ polytope)



For $r=3$ is the answer unique? robust?

[Vavasis] uses this gadget to encode a truth assignment for a variable in 3-SAT

When is NMF easy?

Suppose $\overset{m \times r}{A}\overset{r \times n}{W}$, $M = AW$ is a separable NMF and $\text{rank}(M) = r$

To keep things simple, let's assume anchor words are unique

π: column → its anchor word

Lemma: Under unique anchor words

j is an anchor word $\Leftrightarrow$ $M^j \notin \text{cone}\{M^{j'} \mid j' \neq j\}$

j^{th} row of M

Proof: (sketch) If j is an anchor word then

M^j = nonnegative multiple of W^j

Now consider $K = \text{cone}\{W^{j'}\} \overset{\supseteq \text{ fact transposed}}{\underset{\subseteq \text{ characterization of anchor words}}{=}} \text{cone}\{M^{j'}\}$

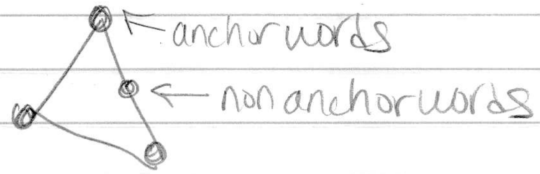
why are these equal?

draw cone for W

= cone for M

where do anchor/non anchor words end up?

Then we have



∎

What would go wrong without unique anchor words?

Anchorwords Algorithm

Remove redundant rows

i.e. multiples of e/o, but keep one per equivalence class

For each row j

if $M^j \notin \text{cone}\{M^{j'} \mid j' \neq j\}$, add to list of anchor words

Set W = anchor words, solve for

$$A = \underset{A \geq 0}{\operatorname{argmin}} \|M - AW\|_F$$

The analysis follows from lemma, notice the answer is unique up to scaling

Can you do faster than solving many LPs?

On HW, you'll show