TD 3 suite 2 - Décomposition en valeurs singulières (SVD)

Polytech Marseille - IRM3
Alexandra Bac

Méthodologie du raisonnement et structures algébriques II

Dans cet exercice, on se place dans un espace **Euclidien**, donc un espace vectoriel de dimension finie mais où l'on a en plus un produit scalaire entre vecteurs (noté $\langle \vec{u}, \vec{v} \rangle$) et donc une norme (donnée par $\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle}$). On a donc à disposition, en plus des ev des TD de Méthodo II, des notions d'orthogonalité, de norme de vecteurs ...

Les propositions suivantes seront admises (mais importantes) :

Proposition 1. *Si A est une matrice carrée symétrique (ie. $A = A^t$), alors elle est forcément diagonalisable en base orthonormale, c'est-à-dire qu'il existe une matrice de changement de base P orthogonale ($P^{-1} = P^t$ ou encore $P^tP = I$) telle que $P^{-1}AP$ soit diagonale.*

Proposition 2. *Si $A = P^{-1}BP$, on dit que les matrices carrées A et B sont semblables (elles représentent la même application linéaire dans des bases différentes) alors elles ont les mêmes valeurs propres.*

Proposition 3 (Produit scalaire, produit matriciel)**.** *Soient X, Y deux vecteurs de $\mathbb{R}^n$, on a :*

(i) $\langle X, Y \rangle = X^t \times Y$ *(on peut donc toujours coder le produit scalaire par un simple produit matriciel)*

(ii) $\|X\|_2^2 = X^t \times X$

Exercice 1 (***)**.** Soit $f : E \to F$ une application linéaire avec $\dim(E) = n$ et $\dim(F) = m$. Soient $\mathcal{B}$ et $\mathcal{B}'$ des bases, respectivement, de E et F et $A = \mathcal{M}_{\mathcal{B},\mathcal{B}'}(f)$ la matrice **rectangulaire** $m \times n$ de f dans ces bases.

(i) Quelle est la taille de $A^t \cdot A$? Montrez que cette matrice est diagonalisable en base orthonormale (on notera $\Delta = \mathrm{diag}(\delta_1, \ldots, \delta_n)$ la matrice diagonale et P la matrice de passage).

(ii) Soit $X_i = (0, \cdots, 0, \overset{i}{1}, 0, \ldots, 0)$. Calculer $X^t \Delta X$ de deux manières différentes pour montrer que pour tout i, on a $\delta_i \geq 0$. On notera donc $\delta_i = \mu_i^2$.

(iii) On va en déduire qu'il existe deux matrices orthogonales U et V telles que $U^tAV = \mathrm{diag}(\mu_1, \ldots, \mu_n)$ (c'est la décomposition en valeurs singulières) :

  (a) Quelle est la taille de AP ? On note f_j la ième colonne de AP, quelle est sa dimension ? Montrer que :
$$f_i^t \times f_j = \mu_i^2 \delta_{ij} \qquad \forall i, j \in \{1 \ldots n\} \tag{1}$$
  où $\delta_{ij} = 0$ si $i \neq j$ et 1 si $i = j$.

(b) Quitte à permuter les vecteurs de la base (donc réordonner P), on peut supposer que les μ_i ont été ordonnés de manière décroissante, donc $\mu_1 \geq \cdots \geq \mu_r$ et $\mu_i = 0$ pour $i \in \{r+1 \dots n\}$ et :

$$\Delta = \left(\begin{array}{ccc|ccc} \mu_1^2 & & & & & \\ & \ddots & & & & \\ & & \mu_r^2 & & & \\ \hline & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{array}\right)$$

Pour $i \in \{1 \dots r\}$, on pose $u_i = {\mu_i}^{-1} \cdot f_i$. Montrez que $\{u_1, \dots, u_r\}$ est une famille orthonormale (ie. les vecteurs sont tous de norme 1 et orthogonaux entre eux).

(c) Montrer que pour $i > r$ on a $\|f_i\|^2 = 0$.

(d) A défaut d'autres vecteurs f_i, on complète $\{u_1, \dots, u_r\}$ en une base orthonormale $\{u_1, \dots, u_m\}$ (c'est toujours faisable et admis ...). On note U la matrice ayant ces vecteurs comme colonnes. Montrer que pour tout $i \in \{1 \dots m\}$ et pour tout $j \in \{1 \dots n\}$, on a : $(U^t AP)_{i,j} = u_i^t \times f_j$.

(e) En déduire que :

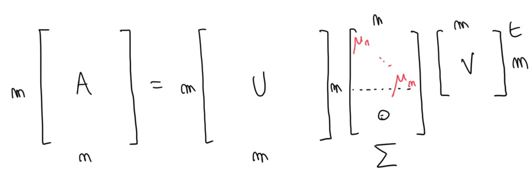
$$(U^t AP)_{i,j} = \mu_i \delta_{i,j}$$

Quelle est par conséquent la matice $U^t AP$?

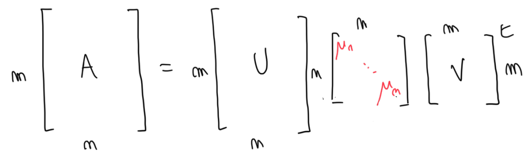
Comme les valeurs propres d'une matrice (en l'occurrence $\mu_1, \dots, \mu_n$, celles de $A^t A$) sont uniques, cette décomposition est également unique. On appelle les μ_i **valeurs singulières de** A (cette décomposition existe toujours, pour toute matrice, contrairement à la diagonalisation).

(iv) En déduire que A est de rang r tout comme $A^t A$ et AA^t.

(v) On obtient donc la décomposition suivante appelée SVD complète :



Montrez que la décomposition suivante (SVD compacte) lui est équivalente :



(vi) On considère la matrice :

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \\ 0 & 1 \end{pmatrix}$$

Calculez la SVD complète de A.

(vii) Montrez que les colonnes de U sont les vecteurs propres orthogonaux de AA^t et que les colonnes de V sont les vecteurs propres orthogonaux de $A^t A$.

Cet outil est fondamental dans de nombreux domaines : informatique graphique, analyse de données, intelligence artificielle ... Voici une application très visuelle d'une propriété de la décomposition SVD. Considérons la SVD compacte :

$$A = U\Sigma V^t \tag{2}$$

où $\Sigma = \operatorname{diag}(\mu_1, \ldots, \mu_n)$. On peut facilement écrire Σ comme une somme de matrices simples. Posons $e_i = (0, \cdots, 0, \overset{i}{1}, 0, \ldots, 0)$, alors $e_i e_i^t$ est la matrice ayant des 0 partout sauf le coefficient (i, i) de la diagonale (qui vaut 1). Par conséquent : $\Sigma = \sum_{i=1}^{n} \mu_i e_i e_i^t = \sum_{i=1}^{r} \mu_i e_i e_i^t$. En injectant cette égalité dans l'équation (2), on obtient donc :

$$A = \sum_{i=1}^{r} \mu_i U e_i e_i^t V^t = \sum_{i=1}^{r} \underbrace{\mu_i (U e_i)(V e_i)^t}_{E_i} \tag{3}$$

En tronquant cette somme, on peut définir :

$$A_k = \sum_{i=1}^{k} E_i$$

Le théorème suivant (théorème d'Eckart-Young) indique qu'on réalise alors une approximation de A par une matrice de rang k et qu'en plus, que c'est la meilleure possible pour ce rang.

Proposition 4 (Théorème d'Eckart-Young)**.** *Soit* $\|A\|_F = \sqrt{\sum_{i,j} A_{i,j}^2}$ *la norme de Frobénius de la matrice* A *(*$\|A - B\|_F$ *caractérise la distance entre* A *et* B *c'est-à-dire l'écart global entre leurs coefficients).*

$$\|A - A_k\|_F = \min_{\substack{B \in \mathcal{M}_{m,n}(\mathbb{R}) \\ \operatorname{rang}(B) = k}} \|A - B\|$$

Or, une image est une matrice !!!! Le résultat est impressionnant.