EECS 127/227AT 9/4/2018

A. Bayen, L. El Ghaoui

Quiz 1: Solutions

1. *permutation matrix.* A matrix $P \in \mathbb{R}^{n\times n}$ is a permutation matrix if it is a permutation of the columns of the $n \times n$ identity matrix.

   (a) For a $n \times n$ matrix A, we consider the products PA and AP. Describe in simple terms what these matrices look like with respect to the original matrix A.

   (b) Show that P is an orthogonal matrix.

   (c) Show that P is invertible, and that $P^{-1} = P^T$.

   Solution:

   (a) Let $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ denote the permutation mapping (which is bijective) corresponding to P. Then we can write $P = \begin{bmatrix} e_{\sigma(1)} & \dots & e_{\sigma(n)} \end{bmatrix}$ where e_i denote the standard basis vectors which are the columns of P.

   We have that PA corresponds to the matrix obtained by permuting the rows of A according to σ. AP corresponds to the matrix obtained by permuting the columns of A according to σ.

   (b) We show $P^\top P = I$. Note $P^\top = \begin{bmatrix} e_{\sigma(1)}^\top \\ \vdots \\ e_{\sigma(n)}^\top \end{bmatrix}$. Thus $(P^\top P)_{ii} = e_{\sigma(i)}^\top e_{\sigma(i)} = 1$ and $(P^\top P)_{ij} = e_{\sigma(i)}^\top e_{\sigma(j)} = 0$ for $i \neq j$[1] so the conclusion follows.

   (c) By definition, P has an orthogonal basis (the standard basis) as its columns. Hence none of its columns can be contained in the span of any of its other columns – so it must be full-rank and invertible. Right-multiplying by P^{-1} in $P^\top P = I$ implies $P^\top = P^{-1}$ since P is orthogonal.

[1] recall $i \neq j \implies \sigma(i) \neq \sigma(j)$.

2. *Rank.* As seen in class, the rank of a $m \times n$ matrix A is the dimension of its range (also called span) $\mathcal{R}(A) := \{Ax \ : \ x \in \mathbb{R}^n\}$.

Figure 1: A black and white image of a painting by Piet Mondrian (1872-1944).

(a) Assume that A takes the form $A = uv^T$, with $u \in \mathbb{R}^m$, $v \in \mathbb{R}^m$. Describe the range of A in simple geometrical terms and find its rank.

(b) Show that the rank of the sum of two matrices of same size A, B is less than the sum of the ranks of A and B. *Hint:* show that $\mathcal{R}(A+B) \subseteq \mathcal{R}(A) + \mathcal{R}(B)$, meaning that any vector in the range of $A+B$ can be expressed as the sum of two vectors, each in the range of A and B respectively. You may also use (without proof) that for any two subspaces S_1, S_2, $\dim(S_1 + S_2) \leq \dim(S_1) + \dim(S_2)$.

(c) Assume that a $m \times n$ matrix A takes the form $A = UV^T$, with $U \in \mathbb{R}^{m \times k}$, $V \in \mathbb{R}^{m \times k}$. Show that the rank of A is less or equal than k. *Hint:* use part (2b).

(d) Consider the image in Fig 1, a gray-scale rendering of a painting by Mondrian (1872-1944). We build a 256×256 matrix A of pixels based on this image by ignoring grey zones (i.e. assigning zero to both light/dark grey zones), assigning $+1$ to horizontal or vertical black lines, $+2$ at the intersections, and zero elsewhere. The horizontal lines occur at row indices 100, 200 and 230, and the vertical ones, at columns indices 50, 230. What is the rank of the matrix?

Solution:

(a) Considering any $x \in \mathbb{R}^m$ we have that $Ax = (v^\top x)u$. Thus the range of A is simply given by the 1-dimensional subspace spanned by u (i.e. the line pointing along u). Evidently, from the above description, the rank of A is 1.

(b) Given any vector $v \in \mathcal{R}(A+B)$, there must exist $x \in \mathbb{R}^m$ such that $v = (A+B)x$. Thus $v = (A+B)x = \underbrace{Ax}_{\in\mathcal{R}(A)} + \underbrace{Bx}_{\in\mathcal{R}(B)}$. So $\mathcal{R}(A+B) \subseteq \mathcal{R}(A) + \mathcal{R}(B)$. From this it follows that $\dim(\mathcal{R}(A+B)) \leq \dim(\mathcal{R}(A)+\mathcal{R}(B))$. Finally, using the second part of the hint, we have that $\dim(\mathcal{R}(A)+\mathcal{R}(B)) \leq \dim(\mathcal{R}(A)) + \dim(\mathcal{R}(B))$ since the range of a matrix is a subspace[2]. Since the rank is the dimension of the range, the conclusion follows.

(c) Let $U = \begin{bmatrix} u_1 & \dots & u_k \end{bmatrix}$ and $V = \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix}$ where u_i and v_i are vectors representing the columns of u and v respectively. We can express the matrix A as,

$$A = UV^\top = \sum_{i=1}^{k} u_i v_i^\top \tag{1}$$

Now using the results of parts (a) and (b) recursively we have that

$$\text{rank}(A) = \text{rank}(\sum_{i=1}^{k} u_i v_i^\top) \leq \text{rank}(\sum_{i=1}^{k-1} u_i v_i^\top) + \underbrace{\text{rank}(u_k v_k^\top)}_{1} \leq \ldots \leq k. \tag{2}$$

as desired.

(d) Assume all the gray regions (both dark and light) correspond to zero entries in the matrix. Based on the description we can express the matrix A as the sum of rank-one matrices corresponding the horizontal/vertical black lines in the image. We once again use e_i to denote the standard basis vectors. Then,

$$\begin{aligned} A = e_{100}(\sum_{i=1}^{50} e_i)^\top + e_{200}(\sum_{i=1}^{256} e_i)^\top + e_{230}(\sum_{i=230}^{256} e_i) \qquad (3) \\ + (\sum_{i=1}^{256} e_i)(e_{50})^\top + (\sum_{i=230}^{256} e_i)(e_{230})^\top \qquad (4) \end{aligned}$$

The first three terms generate the portions of matrix corresponding to the three horizontal black lines (from top to bottom). The last two terms generate the portions of matrix corresponding to the two vertical lines (from left to right). Note the importance of adding +2 at the intersections for the validity of the aforementioned description.

Finally note the set vectors e_{100}, e_{200}, e_{230}, $\sum_{i=1}^{256} e_i$, $\sum_{i=230}^{256} e_i$ (corresponding to u_i's) are linearly independent *and* the set of vectors $\sum_{i=1}^{50} e_i$, $\sum_{i=1}^{256} e_i$, $\sum_{i=230}^{256} e_i$, e_{50}, e_{230} (corresponding to v_i's) are linearly independent. Thus the rank of A is *exactly* the sum of the ranks of each of the rank-one components in the aforementioned decomposition – so $\text{rank}(A) = 5$.

[2]This fact in the hint can be proved by taking a basis of S_1 and extending it to a basis of S_2 (during which we can only add *at most* $\dim(S_2)$ basis vectors). This extended basis must now also be a basis of $S_1 + S_2$. Hence we must have that $\dim(S_1 + S_2) \leq \dim(S_1) + \dim(S_2)$.