MathTools HW 4

1. **The matrix rank.**

   (a) Prove that
   $$\mathrm{rank}(BC) \leq \min\left(\mathrm{rank}(B), \mathrm{rank}(C)\right) .$$

   (b) Prove that for a matrix $A \in M_{n\times m}(\mathbb{F})$, $\mathrm{rank}(A)$ is the smallest k such that one can find $B \in M_{n\times k}(\mathbb{F}), \in M_{k\times m}(\mathbb{F})$ with $A = BC$.

   Hint: There are two directions to show: (i) If there is such a factorization of A, then $\mathrm{rank}(A) \leq k$. (ii) For $k = \mathrm{rank}(A)$, such B and C can be found. One way to proceed is the following: since $\dim(\mathrm{col}(A)) = k$, there is a basis of size k that spans the column space. Let B be the matrix whose columns are these basis vectors. How to choose C?

   (c) Application: let $z_1, \ldots, z_n \in \mathbb{R}^d$ and $x_i = Bz_i \in \mathbb{R}^m$. Define the **Gram matrix** of the points $x_1, \ldots, x_n$: $G_{ij} = \langle x_i, x_j \rangle$; this is an n-by-n matrix. This matrix is particularly interesting when $n, m \gg d$.

   Show that $\mathrm{rank}(G) \leq d$.

   Context: This is a toy model for a technique in ML/statistics (very related to PCA). Suppose $x_1, \ldots, x_n$ are sampled from some data set; one doesn't observe the data points themselves, but instead only has access to pair-wise "distance" or "affinities" (in this case, the correlation $\langle x_i, x_j \rangle$). If one observes that the matrix G (which we *can* form from the given data) is very low-rank, this means that the data set itself has a low-rank latent structure.

2. **QR factorization.** Recall the **Gram-Schmidt procedure**, which takes a basis $v_1, \ldots, v_n \in \mathbb{R}$ and outputs an **orthonormal** basis $u_1, \ldots, u_n$ such that $\mathrm{span}(u_1, \ldots, u_k) = \mathrm{span}(v_1, \ldots, v_k)$ for all $k = 1, \ldots, n$. The procedure builds the u_i-s iteratively: starting from $u_1 = v_1/\|v_1\|$, it constructs
$$\begin{aligned} \tilde{u}_i &= v_i - \sum_{j=1}^{i-1} \langle v_i, u_J \rangle u_j \,, \\ u_i &= \tilde{u}_i / \|\tilde{u}_i\| \,, \end{aligned}$$
sequentially for $i = 2, 3, \ldots$.

   Let $A \in M_{n\times n}(\mathbb{R})$ be a square matrix. Use the above to show that there is an orthogonal matrix Q and an upper triangular R such that $A = QR$.

   Remark: R is upper triangular if all its entries below the diagonal are 0, that is, $R_{ij} = 0$ for $i > j$.

3. **Polynomial interpolation.** Let $x_1, \dots, x_n \in \mathbb{R}$ and $y_1, \dots, y_n \in \mathbb{R}$ be numbers (where the x_i-s are all distinct). Show that there is a polynomial p of degree at most $\le n-1$, such that $p(x_i) = y_i$ for all i.

   Hint: Start by constructing polynomials q_i such that $q_i(x_i) = 1$ and $q_i(x_j) = 0$ for all $j \neq i$.

4. **Vandermonde determinant.** Let $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ be $n \ge 2$ numbers. The corresponding **Vandermonde matrix** is

$$V(\alpha_1, \dots, \alpha_n) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \vdots & \vdots & & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \dots & \alpha_n^{n-1} \end{bmatrix} \in M_{n \times n}(\mathbb{C}),$$

   or in other words, $V_{i,j} = \alpha_j^{i-1}$. Prove that

$$|\det V(\alpha_1, \dots, \alpha_n)| = \prod_{i<j} |\alpha_i - \alpha_j| .$$

   In particular, V is invertible when all the α-s are distinct.

   Hint: Consider the polynomial $p(x) = V(x, \alpha_2, \dots, \alpha_n)$. What are its roots? What is its degree, and what is the coefficient on the leading term? It might be useful (but not necessarily) to remember how to calculate the determinant using an expansion into minors: for any j,

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det(A_{\{-i,-j\}}),$$

   where $A_{\{-i,-j\}}$ is the matrix obtained by erasing the i-th row and j-th column of A.

5. **Similar matrices and change of basis.** Show that if $A \in M_{n\times n}(\mathbb{F})$ and $B \in M_{n\times n}(\mathbb{F})$ are similar matrices, then they represent the same linear mapping with respect to two (possibly, if $A \neq B$) different bases. That is, show that there is a vector space V, linear map $T : V \to V$ and bases $\mathcal{B}, \mathcal{B}'$ such that

$$A = [T]_{\mathcal{B}}^{\mathcal{B}}, \quad B = [T]_{\mathcal{B}'}^{\mathcal{B}'} .$$

   Recall: A, B are similar if there is an invertible P such that $A = PBP^{-1}$.

   Hint: Obviously, there aren't many natural choices for V and T. Take $V = \mathbb{F}^n$ and $T(x) = Ax$.

6. **Norms and metrics.**
   (a) Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms. Show that for $\alpha, \beta > 0$, $\|\cdot\| = \alpha\|\cdot\|_1 + \beta\|\cdot\|_2$ is also a norm.
   (b) Let $d(\cdot,\cdot)$ be a metric. Show that

$$d_*(x,y) = \frac{d(x,y)}{1 + d(x,y)}$$

   is a metric.