Self grades are due at 11 PM on February 2, 2024.

1. **Least Squares**

The Michaelis-Menten model for enzyme kinetics relates the rate y of an enzymatic reaction to the concentration x of a substrate, as follows:

$$y = \frac{\beta_1 x}{\beta_2 + x}, \tag{1}$$

for constants $\beta_1, \beta_2 > 0$.

(a) Show that the model can be expressed as a linear relation between the values $1/y = y^{-1}$ and $1/x = x^{-1}$. Specifically, give an equation of the form $y^{-1} = w_1 + w_2 x^{-1}$, specifying the values of w_1 and w_2 in terms of β_1 and β_2.

Solution: Inverting each side of the equation, we have

$$\begin{aligned} y^{-1} &= \left(\frac{\beta_1 x}{\beta_2 + x}\right)^{-1} && (2)\\ &= \frac{\beta_2 + x}{\beta_1 x} && (3)\\ &= \frac{\beta_2}{\beta_1 x} + \frac{x}{\beta_1 x} && (4)\\ &= \frac{\beta_2}{\beta_1} x^{-1} + \frac{1}{\beta_1} && (5)\\ &= \frac{1}{\beta_1} + \frac{\beta_2}{\beta_1} x^{-1}. && (6)\\ & && (7) \end{aligned}$$

The above equation has exactly the desired form $y^{-1} = w_1 + w_2 x^{-1}$ for $w_1 = \frac{1}{\beta_1}$ and $w_2 = \frac{\beta_2}{\beta_1}$.

(b) In general, reaction parameters β_1 and β_2 (and, thus, w_1 and w_2) are not known a priori and must be fit from data — for example, using least squares. Suppose you collect m measurements (x_i, y_i), $i = 1, \ldots, m$ over the course of a reaction. Formulate the least squares problem

$$\vec{w}^\star = \underset{\vec{w}}{\operatorname{argmin}} \|X\vec{w} - \vec{y}\|_2^2, \tag{8}$$

where $\vec{w}^\star = \begin{bmatrix} w_1^\star & w_2^\star \end{bmatrix}^\top$, and you must specify $X \in \mathbb{R}^{m\times 2}$ and $\vec{y} \in \mathbb{R}^m$. Specifically, your solution should include explicit expressions for X and $\vec{y}$ as a function of (x_i, y_i) values and a final expression for $\vec{w}^\star$ in terms of X and $\vec{y}$, which should contain only matrix multiplications, transposes, and inverses.

Assume without loss of generality that $x_1 \neq x_2$.

Solution: To formulate the least squares problem as stated, X and $\vec{y}$ values should be set to

$$X = \begin{bmatrix} 1 & \cdots & 1 \\ x_1^{-1} & \cdots & x_m^{-1} \end{bmatrix}^\top, \quad \vec{y} = \begin{bmatrix} y_1^{-1} & \cdots & y_m^{-1} \end{bmatrix}^\top. \tag{9}$$

To solve this least squares problem, we note that the optimal residual vector $X\vec{w}^\star - \vec{y}$ must be orthogonal to $\mathcal{R}(X)$ by the orthogonality principle, and we have

$$X^\top (X\vec{w}^\star - \vec{y}) = 0. \tag{10}$$

Rearranging we get,

$$\vec{w}^\star = (X^\top X)^{-1} X^\top \vec{y}. \tag{11}$$

(c) Assume that we have used the above procedure to calculate values for $w_1^\star$ and $w_2^\star$, and we now want to estimate $\widehat{\vec{\beta}} = \begin{bmatrix} \widehat{\beta}_1 & \widehat{\beta}_2 \end{bmatrix}^\top$. Write an expression for $\widehat{\vec{\beta}}$ in terms of $w_1^\star$ and $w_2^\star$.

Solution: To calculate $\widehat{\vec{\beta}}$, we can simply reverse the calculations from part (a):

$$w_1 = \frac{1}{\beta_1} \implies \beta_1 = \frac{1}{w_1}, \tag{12}$$

$$w_2 = \frac{\beta_2}{\beta_1} \implies \beta_2 = \beta_1 w_2 = \frac{w_2}{w_1}. \tag{13}$$

Thus, $\widehat{\vec{\beta}} = \begin{bmatrix} \frac{1}{w_1^\star} & \frac{w_2^\star}{w_1^\star} \end{bmatrix}^\top$.

NOTE: This problem was taken (with some edits) from the textbook *Optimization Models* by Calafiore and El Ghaoui.

2. Subspaces and Dimensions

Consider the set $\mathcal{S}$ of points $(x_1, x_2, x_3) \in \mathbb{R}^3$ such that

$$x_1 + 2x_2 + 3x_3 = 0, \quad 3x_1 + 2x_2 + x_3 = 0. \tag{14}$$

(a) Find a 2×3 matrix A for which $\mathcal{S}$ is exactly the null space of A.

Solution: Recall the definition of the null space of a matrix A as the set of all vectors $\vec{x}$ such that $A\vec{x} = \vec{0}$. The equations

$$x_1 + 2x_2 + 3x_3 = 0 \tag{15}$$

$$3x_1 + 2x_2 + x_3 = 0 \tag{16}$$

can be written in matrix-vector form as

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{17}$$

The set of $\vec{x} = \begin{bmatrix} x_1, x_2, x_3 \end{bmatrix}^\top$ which satisfy this equation form the null space of $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$. This is the matrix we are looking for.

(b) Determine the dimension of $\mathcal{S}$ and find a basis for it.

Solution: Recall the definitions of a basis and the dimension of a subspace, which are related. A basis for a space is a set of linearly independent vectors that span the space. The dimension of this space is then the number of vectors in the basis.

To find the dimension, we solve the equation and find that any solution to the equations is of the form $x_1 = x_3$, $x_2 = -2x_3$, where x_3 is free. Thus, the solutions are of the form $\begin{bmatrix} 1, -2, 1 \end{bmatrix}^\top u$ for $u \in \mathbb{R}$, and so $\mathcal{S} = \operatorname{span}\left(\begin{bmatrix} 1, -2, 1 \end{bmatrix}^\top\right)$. Hence, the dimension of $\mathcal{S}$ is 1, and a basis for $\mathcal{S}$ is the vector $\begin{bmatrix} 1, -2, 1 \end{bmatrix}^\top$.

3. Vector Spaces and Rank

The *rank* of a $m \times n$ matrix A, $\mathrm{rank}(A)$, is the dimension of its *range*, also called span, and denoted $\mathcal{R}(A) := \{A\vec{x} \,:\, \vec{x} \in \mathbb{R}^n\}$.

(a) Assume that $A \in \mathbb{R}^{m\times n}$ takes the form $A = \vec{u}\vec{v}^\top$, with $\vec{u} \in \mathbb{R}^m$, $\vec{v} \in \mathbb{R}^n$, and $\vec{u}, \vec{v} \neq \vec{0}$. (Note that a matrix of this form is known as a *dyad*.) Find the rank of A.

HINT: Consider the quantity $A\vec{x}$ for arbitrary $\vec{x}$, i.e., what happens when you multiply any vector by matrix A.

Solution: For any $\vec{x} \in \mathbb{R}^n$, we have that $A\vec{x} = \vec{u}\vec{v}^\top\vec{x} = \vec{u}(\vec{v}^\top\vec{x}) = (\vec{v}^\top\vec{x})\vec{u}$. Note that $\vec{v}^\top\vec{x}$ is a scalar that can take on any value depending on choice of $\vec{x}$. Since the range of A is the subspace reachable through any choice of $\vec{x}$, $\mathcal{R}(A)$ is simply the 1-dimensional subspace spanned by $\vec{u}$ (i.e., the line pointing along $\vec{u}$). Since a single vector (namely, $\vec{u}$) spans $\mathcal{R}(A)$, the rank of A is 1.

(b) Show that for arbitrary $A, B \in \mathbb{R}^{m\times n}$,

$$\mathrm{rank}(A + B) \leq \mathrm{rank}(A) + \mathrm{rank}(B), \tag{18}$$

i.e., the rank of the sum of two matrices is less than or equal to the sum of their ranks.

HINT: First, 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. Remember that for any matrix A, $\mathcal{R}(A)$ is a subspace, and for any two subspaces S_1 and S_2, $\dim(S_1 + S_2) \leq \dim(S_1) + \dim(S_2)$.[1] (Note that the sum of vector spaces $S_1 + S_2$ is not equivalent to $S_1 \cup S_2$, but is defined as $S_1 + S_2 := \{\vec{s_1} + \vec{s_2} | \vec{s_1} \in S_1, \vec{s_2} \in S_2\}$.)

Solution: Given any vector $\vec{v} \in \mathcal{R}(A + B)$, there must by definition exist $\vec{x} \in \mathbb{R}^n$ such that $\vec{v} = (A+B)\vec{x}$. Thus, $\vec{v} = (A + B)\vec{x} = \underbrace{A\vec{x}}_{\in\mathcal{R}(A)} + \underbrace{B\vec{x}}_{\in\mathcal{R}(B)}$, so $\mathcal{R}(A + B) \subseteq \mathcal{R}(A) + \mathcal{R}(B)$, as hinted.

Computing the dimension of each side of the subset relationship, it follows that

$$\dim(\mathcal{R}(A + B)) \leq \dim(\mathcal{R}(A) + \mathcal{R}(B)). \tag{19}$$

Using the second part of the hint, we have that

$$\dim(\mathcal{R}(A) + \mathcal{R}(B)) \leq \dim(\mathcal{R}(A)) + \dim(\mathcal{R}(B)). \tag{20}$$

Combining the previous two equations,

$$\dim(\mathcal{R}(A + B)) \leq \dim(\mathcal{R}(A)) + \dim(\mathcal{R}(B)), \tag{21}$$

i.e., by definition,

$$\mathrm{rank}(A + B) \leq \mathrm{rank}(A) + \mathrm{rank}(B) \tag{22}$$

as desired.

(c) Consider an $m \times n$ matrix A that takes the form $A = UV^\top$, with $U \in \mathbb{R}^{m\times k}$, $V \in \mathbb{R}^{n\times k}$. Show that the rank of A is less or equal than k. *HINT: Use parts (a) and (b), and remember that this decomposition can*

[1] This fact 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 $S_1 + S_2$. Thus, $\dim(S_1 + S_2) \leq \dim(S_1) + \dim(S_2)$.

also be written as the dyadic expansion

$$A = UV^\top = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_k \end{bmatrix} \begin{bmatrix} \vec{v}_1^\top \\ \vdots \\ \vec{v}_k^\top \end{bmatrix} = \sum_{i=1}^{k} \vec{u}_i \vec{v}_i^\top, \tag{23}$$

for $U = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_k \end{bmatrix}$ *and* $V = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_k \end{bmatrix}$.

Solution: Starting with the dyadic expansion above, iteratively pulling out terms from this summation, and using the result from (a) that the rank of a dyadic matrix is 1 (or 0, if any $\vec{v}_i = \vec{0}$), we know by the rank relation from (b) that

$$\operatorname{rank}(A) = \operatorname{rank}\left(\sum_{i=1}^{k} \vec{u}_i \vec{v}_i^\top\right) \le \operatorname{rank}\left(\sum_{i=1}^{k-1} \vec{u}_i \vec{v}_i^\top\right) + \underbrace{\operatorname{rank}\left(\vec{u}_k \vec{v}_k^\top\right)}_{\text{0 or 1}} \le \dots \le k, \tag{24}$$

as desired.

4. Norms

(a) Show that the following inequalities hold for any vector $\vec{x} \in \mathbb{R}^n$:

$$\frac{1}{\sqrt{n}} \|\vec{x}\|_2 \leq \|\vec{x}\|_\infty \leq \|\vec{x}\|_2 \leq \|\vec{x}\|_1 \leq \sqrt{n}\,\|\vec{x}\|_2 \leq n\,\|\vec{x}\|_\infty . \tag{25}$$

NOTE: We can interpret different norms as different ways of computing distance between two points $\vec{x}, \vec{y} \in \mathbb{R}^2$. The ℓ^2 norm is the distance as the crow flies (i.e. point-to-point distance), the ℓ^1 norm, also known as the Manhattan distance is the distance you would have to cover if you were to navigate from $\vec{x}$ to $\vec{y}$ via a rectangular street grid, and the ℓ^∞ norm is the maximum distance that you have to travel in either the north-south or the east-west direction.

Solution: We have

$$\|\vec{x}\|_2^2 = \sum_{i=1}^n x_i^2 \leq n \cdot \max_i\, x_i^2 = n \cdot \|\vec{x}\|_\infty^2 . \tag{26}$$

Also, $\|\vec{x}\|_\infty \leq \sqrt{x_1^2 + \ldots + x_n^2} = \|\vec{x}\|_2$. The inequality $\|\vec{x}\|_2 \leq \|\vec{x}\|_1$ is obtained after squaring both sides, and checking that

$$\sum_{i=1}^n x_i^2 \leq \sum_{i=1}^n x_i^2 + \sum_{i \neq j} |x_i x_j| = \left(\sum_{i=1}^n |x_i| \right)^2 = \|\vec{x}\|_1^2 . \tag{27}$$

The condition $\|\vec{x}\|_1 \leq \sqrt{n}\,\|\vec{x}\|_2$ is due to the Cauchy-Schwarz inequality

$$\left| \vec{z}^\top \vec{y} \right| \leq \|\vec{y}\|_2 \cdot \|\vec{z}\|_2 , \tag{28}$$

applied to the two vectors $y = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}^\top$ and $\vec{z} = |\vec{x}| = \begin{bmatrix} |x_1| & \cdots & |x_n| \end{bmatrix}$.

Finally, $\sqrt{n}\,\|\vec{x}\|_2 \leq n\,\|\vec{x}\|_\infty$, is achieved by an algebraic manipulation of the first derived bound using the fact that $\sqrt{n} = \frac{n}{\sqrt{n}}$.

(b) We define the *sparsity* of the vector $\vec{x}$ as the number of non-zero elements in $\vec{x}$. This is also commonly known as the ℓ^0 norm of the vector $\vec{x}$, denoted by $\|\vec{x}\|_0$. Show that for any non-zero vector x,

$$\|\vec{x}\|_0 \geq \frac{\|\vec{x}\|_1^2}{\|\vec{x}\|_2^2} . \tag{29}$$

Find all vectors $\vec{x}$ for which the lower bound is attained.

Solution: Let us apply the Cauchy-Schwarz inequality with $\vec{z} = |\vec{x}|$ again, and with $\vec{y}$ a vector with $y_i = 1$ if $x_i \neq 0$, and $y_i = 0$ otherwise. We have $\|\vec{y}\|_2 = \sqrt{k}$, with $k = \|\vec{x}\|_0$. Hence

$$\left| \vec{z}^\top \vec{y} \right| = \|\vec{x}\|_1 \leq \|\vec{y}\|_2 \cdot \|\vec{z}\|_2 = \sqrt{k} \cdot \|\vec{x}\|_2 , \tag{30}$$

which proves the result. The bound is attained for vectors with k non-zero elements, all with the same magnitude.

5. Homework Process

With whom did you work on this homework? List the names and SIDs of your group members.

NOTE: If you didn't work with anyone, you can put "none" as your answer.