最优化方法：基础知识习题

2024 年 3 月 3 日

1. 给定线性方程组

$$x_1 + x_2 + 2x_3 + x_4 = 1$$
$$x_1 - 2x_2 - x_4 = -2$$

请求取该方程组的解。

2. 函数 $\langle \cdot, \cdot \rangle_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ 定义为 $\langle \mathbf{x}, \mathbf{y} \rangle_2 = 2x_1y_1 + 3x_2y_1 + 3x_1y_2 + 5x_2y_2$, 其中, $\mathbf{x} = [x_1, x_2]^\top$, $\mathbf{y} = [y_1, y_2]^+$ 。试证明 $\langle \cdot, \cdot \rangle_2$ 满足内积的 3 个性质。

3. 证明下述定理成立：对称矩阵 $\mathbf{Q}$ 是正定（半正定）的，当且仅当 $\mathbf{Q}$ 的所有特征值是正的（非负的）

4. 给定以下条件, 试寻找从 $\{e_1, e_2, e_3\}$ 到 $\{e_1', e_2', e_3'\}$ 的变换矩阵 $\mathbf{T}$:

   a. $e_1' = e_1 + 3e_2 - 4e_3, e_2' = 2e_1 - e_2 + 5e_3, e_3' = 4e_1 + 5e_2 + 3e_3$

   b. $e_1 = e_1' + e_2' + 3e_3', e_2 = 2e_1' - e_2' + 4e_3' \cdot e_3 = 3e_1' + 5e_3'$

5. 考虑 $\mathbb{R}^3$ 的两组基 $\{e_1, e_2, e_3\}$ 和 $\{e_1', e_2', e_3'\}$, 其中 $e_1 = 2e_1' + e_2' - e_3', e_2 = 2e_1' - e_2' + 2e_3', e_3 = 3e_1' + e'$。如果在基 $\{e_1, e_2, e_3\}$ 下, 某一线性变换的矩阵表示为

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

试确定该线性变换在基 $\{e_1', e_2', e_3'\}$ 下的矩阵表示。

6. 试确定矩阵

$$\mathbf{A} = \begin{bmatrix} 4 & -2 & 0 \\ 2 & 1 & -1 \\ 2 & -3 & 1 \end{bmatrix}$$

的零空间。

7. 试判定二次型

$$\mathbf{x}^{\top}\left[\begin{array}{cc}1 & -8 \\ 1 & 1\end{array}\right]\mathbf{x}$$

是正定的，半正定的，负定的半负定的还是不定的?

8. 给定如下 3 种二次型, 试判断它们是正定的、半正定的、负定的、半负定的还是不定的?

a. $f\left(x_1, x_2, x_3\right)=x_2^2$

b. $f\left(x_1, x_2, x_3\right)=x_1^2+2 x_2^2-x_1 x_3$

c. $f\left(x_1, x_2, x_3\right)=x_1^2+x_3^2+2 x_1 x_2+2 x_1 x_3+2 x_2 x_3$

9. 试确定一个线性变换, 使得如下二次型能够变换为对角阵:

$$f\left(x_1, x_2, x_3\right)=4 x_1^2+x_2^2+9 x_3^2-4 x_1 x_2-6 x_2 x_3+12 x_1 x_3$$

10. 考虑二次型:

$$f\left(x_1, x_2, x_3\right)=x_1^2+x_2^2+5 x_3^2+2 \xi x_1 x_2-2 x_1 x_3+4 x_2 x_3$$

试确定参数 ξ 的取值范围, 使得该二次型是正定的。

11. 考虑函数 $\langle\cdot, \cdot\rangle_Q: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, 定义为 $\langle\mathbf{x}, \mathbf{y}\rangle_Q=\mathbf{x}^{\top} \mathbf{Q} \mathbf{y}$, 其中, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, 且 $\mathbf{Q} \in \mathbb{R}^{n \times n}$ 是对称正定阵。试证明 $\langle\cdot, \cdot\rangle_Q$ 满足内积的 3 个条件。

12. 考虑 $\mathbb{R}^n$ 上的向量范数 $\|\cdot\|_{\infty}$ ，定义为 $\|x\|_{\infty}=\max_i\left|x_i\right|$ ，其中 $x=\left[x_1, \cdots, x_n\right]^{\top}$ 。类似地, 可以定义 $\mathbb{R}^m$ 上的向量范数 $\|\cdot\|_{\infty}$ 。试证明, 由这些向量范数导出的矩阵范数为

$$\|\mathbf{A}\|_{\infty}=\max_i \sum_{k=1}^n\left|a_{i k}\right|$$

其中, $a_{i j}$ 表示 $\mathbf{A} \in \mathbb{R}^{m \times n}$ 的第 (i, j) 个元素。

13. 设 $f(\mathbf{x})=\|\mathbf{A x}-\mathbf{y}\|^2$, 其中 $\mathbf{A}$ 为列满秩矩阵.

a. 使用逐分量计算梯度方式计算 $\frac{\partial f}{\partial x_k}$;

b. 使用整体的方法计算梯度 $\frac{\partial f}{\partial x}$.

14. 设 $f(\mathbf{x})=tr(\sigma^2 \mathbf{I}+\mathbf{x x}^{\top}), \sigma>0, \mathbf{I}$ 为单位阵,$\mathbf{x} \in \mathbb{R}^n$.

a. 使用逐分量计算梯度方式计算 $\frac{\partial f}{\partial x_k}$;

b. 使用整体的方法计算梯度 $\frac{\partial f}{\partial x}$.

15. 设 $f(\mathbf{x})=\frac{1}{4}\|\mathbf{A}-\mathbf{x x}^{\top}\|^2$, 其中 $\mathbf{A}$ 为对称正定矩阵,$\mathbf{x} \in \mathbb{R}^n$.

a. 使用逐分量计算梯度方式计算 $\frac{\partial f}{\partial x_k}$;

b. 使用整体的方法计算梯度 $\frac{\partial f}{\partial \mathbf{x}}$.(提示:$tr(\mathbf{A}^\top \mathbf{A}) = \|\mathbf{A}\|^2$)

16. 设 $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{Q}\mathbf{x}, \mathbf{x} \in \mathbb{R}^n$ 其中 $\mathbf{Q} \in \mathbb{R}^{n\times n}$ 为正定矩阵.

a. 使用逐分量计算梯度方式计算 $\frac{\partial f}{\partial q_{ij}}$;

b. 使用整体的方法计算梯度 $\frac{\partial f}{\partial \mathbf{Q}}$.

17. 设 $f(\mathbf{X}) = tr(\mathbf{A}\mathbf{X}\mathbf{B}), \mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{X} \in \mathbb{R}^{n\times p}, \mathbf{B} \in \mathbb{R}^{p\times m}$, 计算梯度 $\frac{\partial f}{\partial \mathbf{X}}$

18. 设 $f(\mathbf{x}) = \sin\log(1 + \mathbf{x}^\top \mathbf{x}), \mathbf{x} \in \mathbb{R}^n$, 计算梯度 $\frac{\partial f}{\partial \mathbf{x}}$.

19. 考虑二次型 $f(\mathbf{x}) = \mathbf{x}^\top \mathbf{Q}\mathbf{x}$, 其中 $\mathbf{x} \in \mathbb{R}^n$,$\mathbf{Q} \in \mathbb{R}^{n\times n}$,$\mathbf{Q}$ 的特征值分解为 $\mathbf{Q} = \mathbf{U}\Lambda\mathbf{U}^\top$, 其中 $\mathbf{U}$ 为正交矩阵, 特征值为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$.

a. 证明

$$\lambda_1 = \max_{\|\mathbf{x}\|=1} \mathbf{x}^\top \mathbf{Q}\mathbf{x} = \max_{\mathbf{x}\neq 0} \frac{\mathbf{x}^\top \mathbf{Q}\mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$$

b. 对于 $\forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq 0$, 都有

$$\lambda_n \|\mathbf{x}\|^2 \leq \mathbf{x}^T \mathbf{Q}\mathbf{x} \leq \lambda_1 \|\mathbf{x}\|^2$$

20. 给定数据集 $\{\mathbf{x}_i\}_{i=1}^N, \mathbf{x}_i \in \mathbb{R}^n$, 设其均值为 $\bar{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^N \mathbf{x}_i$, 设以 $\mathbf{w}$ 为方向向量, 过点 $\bar{\mathbf{x}}$ 的直线为

$$\mathbf{L} = \{\mathbf{y} \in \mathbb{R}^n | \mathbf{y} = \bar{\mathbf{x}} + t\mathbf{w}, t \in \mathbb{R}\}$$

对任何 $\mathbf{x} \in \mathbf{R}^n$, 定义 $\tilde{\mathbf{x}}_i = \arg\min_{\mathbf{y}\in\mathbf{L}} \|\mathbf{x}_i - \mathbf{y}\|$ 为 $\mathbf{x}$ 在 $\mathbf{L}$ 上的投影. 请计算数据 $\mathbf{x}_i$ 在 $\mathbf{L}$ 的投影所对应的点 $\tilde{\mathbf{x}}$

21. 完成以下矩阵求导公式的推导:

a.

$$\frac{\partial \det(\mathbf{X})}{\partial \mathbf{X}} = \det(\mathbf{X})\mathbf{X}^{-\top}$$

b.

$$\frac{\partial \mathbf{Y}^{-1}}{\partial x} = -\mathbf{Y}^{-1}\frac{\partial \mathbf{Y}}{\partial x}\mathbf{Y}^{-1}$$

22. 证明：所有由 $m \times n$ 矩阵组成的集合 $\mathbb{R}^{m\times n}$，在矩阵加法和数值-矩阵乘法下构成线性空间。

23. 证明：所有定义在 $[a, b]$ 上的连续函数构成的集合

$$\mathcal{C} = \{f | f : [a, b] \to \mathbb{R}, f \text{ is continuous}\}$$

在 f,$g \in \mathcal{C}$ 和实数 α，有 $f + g \in \mathcal{C}, \alpha f \in \mathcal{C}$ 两个运算下构成线性空间。

24. 证明：任意过原点的直线都构成一个子空间.

25. 证明：对于内积空间 $(\mathbb{V}, \langle \cdot, \cdot \rangle)$，定义 $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ 此定义满足范数 3 个性质。

26. 证明： $f : \mathbb{R}^n \to \mathbb{R}, f(\mathbf{x}) = \mathbf{a}^\top \mathbf{x}$，映射 f 为线性映射。

27. 证明：令 $tr : \mathbb{R}^{n\times n} \to \mathbb{R}$ 为矩阵迹，$tr(\cdot)$ 是一个矩阵集上的线性映射。

28. 证明：对于 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, 有 $\mathbf{x}^\top \mathbf{y} = tr(\mathbf{x}^\top \mathbf{y}) = tr(\mathbf{x}\mathbf{y}^\top)$

29. 证明：对于 $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ 和矩阵 $\mathbf{A} \in \mathbb{R}^{n\times n}$，有

$$\mathbf{x}^\top \mathbf{A}\mathbf{y} = tr(\mathbf{x}^\top \mathbf{A}\mathbf{y}) = tr(\mathbf{y}\mathbf{x}^\top \mathbf{A}) = tr(\mathbf{A}\mathbf{y}\mathbf{x}^\top)$$

30. 证明：谱范数满足范数的 3 个性质：

$$\|\mathbf{A}\| = \max_{x\neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|_2$$

31. 证明下述式子成立：

设 $\mathbf{A}$ 为对称矩阵，$\lambda_{min}, \lambda_{max}$ 分别为其最小和最大特征值，则成立

$$\lambda_{min}(\mathbf{A}) = \min_{\mathbf{x}^\top \mathbf{x}=1} \mathbf{x}^\top \mathbf{A}\mathbf{x}, \lambda_{max}(\mathbf{A}) = \max_{\mathbf{x}^\top \mathbf{x}=1} \mathbf{x}^\top \mathbf{A}\mathbf{x}$$

上述结论表明，函数 $\mathbf{x}^\top \mathbf{A}\mathbf{x}$ 在单位球 $\|\mathbf{x}\| = 1$ 上的取值范围总是落在区间 $[\lambda_{min}, \lambda_{max}]$ 的范围内上述结论可以等价的写成

$$\lambda_{min}(\mathbf{A}) = \min_{\mathbf{x}^\top \mathbf{x}=1} \mathbf{x}^\top \mathbf{A}\mathbf{x} = \min_{\mathbf{x}\neq 0} \frac{\mathbf{x}^\top \mathbf{A}\mathbf{x}}{\mathbf{x}^\top \mathbf{x}},$$

$$\lambda_{max}(\mathbf{A}) = \max_{\mathbf{x}^\top \mathbf{x}=1} \mathbf{x}^\top \mathbf{A}\mathbf{x} = \max_{\mathbf{x}\neq 0} \frac{\mathbf{x}^\top \mathbf{A}\mathbf{x}}{\mathbf{x}^\top \mathbf{x}},$$

32. 证明：正定矩阵都是可逆的, 并且 $\mathbf{Q}^{-1} = \mathbf{U}\Lambda^{-1}\mathbf{U}^\top$. 则，正定矩阵的逆矩阵也是正定矩阵。

33. 二次型 $f(\mathbf{x}) = 3{x_1}^2 + 4{x_2}^2 + 3{x_3}^2 + 2x_1x_3$

a. 将此二次型写成矩阵-向量形式，并求该二次型的标准型.

b. 证明：$\min_x \frac{f(\mathbf{x})}{\mathbf{x}^\top \mathbf{x}} = 2$.

c. 证明 $\{\mathbf{x} \in \mathbb{R}^3 | f(\mathbf{x}) = 1\}$ 是一个椭圆，并计算椭圆的长轴与短轴之比.

d. 计算 $\nabla^2 f(\mathbf{x})$ 的条件数.

34. 证明下述三个等式

a. $\mathrm{d}tr(\mathbf{X}) = tr(\mathrm{d}\mathbf{X})$, 因此 $\frac{\partial tr(\mathbf{X})}{\partial \mathbf{X}} = \mathbf{I}$

b. 乘法法则：$\mathrm{d}(\mathbf{XY}) = (\mathrm{d}\mathbf{X})\mathbf{Y} + \mathbf{X}(\mathrm{d}\mathbf{Y})$

c. $\mathrm{d}(\mathbf{X}^\top) = (\mathrm{d}\mathbf{X})^\top$

35. $f(\mathbf{X}) = tr(\mathbf{X}^\top \mathbf{A}\mathbf{X})$, 计算 $\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}}$

36. $f(\mathbf{X}) = \log(\det \mathbf{X})$, 其中 $\mathbf{X}$ 为正定矩阵, 计算 $\frac{\partial f(\mathbf{X})}{\partial \mathbf{X}}$

37. 最小二乘问题：$f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$, 其中 $\mathbf{A}$ 是满秩矩阵，计算其梯度和 Hessian 矩阵.

38.log-sum-exp 函数：

$$f(\mathbf{x}) = \log \sum_{i=1}^{n} \exp x_i$$

计算其梯度和 Hessian 矩阵，证明 Hessian 矩阵正定.