数值分析内部课程习题解答

第八章

SCAU DataHub

目录

1 Sparsity and structure

Solution 1.1. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-1-sparsity-and-structure.ipynb`.

Solution 1.2. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-1-sparsity-and-structure.ipynb`.

Solution 1.3. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-1-sparsity-and-structure.ipynb`.

Solution 1.4. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-1-sparsity-and-structure.ipynb`.

Solution 1.5. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-1-sparsity-and-structure.ipynb`.

Solution 1.6. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-1-sparsity-and-structure.ipynb`.

2 Power iteration

Solution 2.1. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-2-power-iteration.ipynb`.

Solution 2.2. 给定

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \mathbf{x}^{(1)} = \begin{bmatrix} 0.4 \\ 0.7 \end{bmatrix}.$$

注意到 $\mathbf{A}$ 的作用是交换向量的两个分量，因此

$$\mathbf{A}\mathbf{x}^{(1)} = \begin{bmatrix} 0.7 \\ 0.4 \end{bmatrix}, \qquad \mathbf{A}^2 = \mathbf{I} \ \Rightarrow\ \mathbf{A}^{2k}\mathbf{x}^{(1)} = \mathbf{x}^{(1)},\ \mathbf{A}^{2k+1}\mathbf{x}^{(1)} = \mathbf{A}\mathbf{x}^{(1)}.$$

也就是说，未归一化的幂迭代会在 $\mathbf{x}^{(1)}$ 与 $\mathbf{A}\mathbf{x}^{(1)}$ 之间来回跳动，一般不会收敛到某个特征向量（除非初值本身就是特征向量，例如 $\mathbf{x}^{(1)} \propto [1,1]^T$ 或 $\mathbf{x}^{(1)} \propto [1,-1]^T$）。

从 (8.2.2) 的角度也能看清原因：$\mathbf{A}$ 的特征值为 $\lambda_1 = 1, \lambda_2 = -1$，满足 $|\lambda_1| = |\lambda_2|$，因此

$$\left|\frac{\lambda_2}{\lambda_1}\right| = 1,$$

对应 (8.2.2) 中的 $\left(\lambda_2/\lambda_1\right)^k$ 项不会衰减到 0，幂迭代不具备“唯一主特征值主导”的收敛机制。

若按照算法 8.2.2 / Function 8.2.3 的归一化规则（用 $y = \mathbf{A}x$ 的最大分量做缩放），也会出现周期 2 的振荡：第一次迭代

$$\mathbf{y}^{(1)} = \mathbf{A}\mathbf{x}^{(1)} = \begin{bmatrix} 0.7 \\ 0.4 \end{bmatrix}, \quad m = 1, \quad \beta_1 = \frac{y_m^{(1)}}{x_m^{(1)}} = \frac{0.7}{0.4} = \frac{7}{4}, \quad \mathbf{x}^{(2)} = \frac{\mathbf{y}^{(1)}}{y_m^{(1)}} = \begin{bmatrix} 1 \\ \frac{4}{7} \end{bmatrix}.$$

第二次迭代

$$\mathbf{y}^{(2)} = \mathbf{A}\mathbf{x}^{(2)} = \begin{bmatrix} \frac{4}{7} \\ 1 \end{bmatrix}, \quad m = 2, \quad \beta_2 = \frac{y_m^{(2)}}{x_m^{(2)}} = \frac{1}{4/7} = \frac{7}{4}, \quad \mathbf{x}^{(3)} = \frac{\mathbf{y}^{(2)}}{y_m^{(2)}} = \begin{bmatrix} \frac{4}{7} \\ 1 \end{bmatrix}.$$

随后 $\mathbf{x}^{(k)}$ 在 $\left[1, \frac{4}{7}\right]^T$ 与 $\left[\frac{4}{7}, 1\right]^T$ 之间周期性切换，β_k 恒等于 7/4，并不会收敛到 ±1。

Solution 2.3. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-2-power-iteration.ipynb`.

Solution 2.4. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-2-power-iteration.ipynb`.

Solution 2.5. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-2-power-iteration.ipynb`.

3 Inverse iteration

Solution 3.1. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-3-inverse-iteration.ipynb`.

Solution 3.2. 给定

$$\mathbf{A} = \begin{bmatrix} 1.1 & 1 \\ 0 & 2.1 \end{bmatrix}, \qquad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

先注意到

$$\mathbf{A}\mathbf{x}_1 = \begin{bmatrix} 2.1 \\ 2.1 \end{bmatrix} = 2.1\,\mathbf{x}_1,$$

因此 $\mathbf{x}_1$ 是 $\lambda = 2.1$ 的特征向量。

对任意移位 $s \neq 2.1$，(8.3.3) 的第一步需要解线性方程

$$(\mathbf{A} - s\mathbf{I})\mathbf{y}_1 = \mathbf{x}_1.$$

由于 $(\mathbf{A} - s\mathbf{I})\mathbf{x}_1 = (2.1 - s)\mathbf{x}_1$，可得

$$\mathbf{y}_1 = \frac{1}{2.1 - s}\,\mathbf{x}_1.$$

随后用无穷范数（或最大分量）归一化 $\mathbf{y}_1$ 不会改变方向，因此一步迭代后的向量满足

$$\mathbf{x}_2 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

对 $s = 1, 2, 1.6$ 均成立。

Solution 3.3. 令

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

它满足 $\mathbf{A}^2 = -\mathbf{I}$，因此特征值为 $\lambda = \pm i$，对应特征向量均为复向量（在实数域中不存在特征向量）。无移位逆迭代相当于反复作用 $\mathbf{A}^{-1}$；但 $\mathbf{A}^{-1} = -\mathbf{A}$ 仍是一个旋转矩阵，迭代会产生周期性旋转而不是收敛。

从收敛判据 (8.3.4) 的角度看也一样：当 $s = 0$ 时，

$$|\lambda_+ - s| = |i| = 1, \qquad |\lambda_- - s| = |-i| = 1,$$

两者到移位点的距离相同，不存在"唯一最近的特征值"，因此逆迭代没有稳定的收敛方向。

取移位 $s = -1$ 也不能改善：此时

$$|\lambda_\pm - s| = |\pm i + 1| = \sqrt{2},$$

仍然等距，(8.3.4) 的主导机制依旧缺失，所以算法仍不会收敛到某个特征向量。

Solution 3.4. 先验证

$$\mathbf{A}_0 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

在 $\lambda = 0$ 处有特征向量。由 $\mathbf{A}_0\mathbf{v} = 0$ 得到 $v_1 + v_2 = 0$，因此可取

$$\mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \qquad \mathbf{A}_0\mathbf{v} = 0.$$

现在考虑轻微扰动

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & \varepsilon \end{bmatrix}, \qquad \mathbf{x}_k = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

并按 (8.3.3)（取零移位）解

$$\mathbf{A}\mathbf{y}_k = \mathbf{x}_k.$$

第二行给出 $\varepsilon\, y_{k,2} = 1$，所以 $y_{k,2} = 1/\varepsilon$；代回第一行得

$$y_{k,1} = 1 - \frac{1}{\varepsilon} = \frac{\varepsilon - 1}{\varepsilon}.$$

因此

$$\mathbf{y}_k = \frac{1}{\varepsilon}\begin{bmatrix} \varepsilon - 1 \\ 1 \end{bmatrix} = \frac{1}{\varepsilon}\begin{bmatrix} -1 + \varepsilon \\ 1 \end{bmatrix}.$$

当 $|\varepsilon|$ 很小时，$\|\mathbf{y}_k\|_\infty = \max\{|y_{k,1}|, |y_{k,2}|\} = 1/|\varepsilon|$，故无穷范数归一化后

$$\frac{\mathbf{y}_k}{\|\mathbf{y}_k\|_\infty} = \begin{bmatrix} -1 + \varepsilon \\ 1 \end{bmatrix} = \mathbf{v} + \begin{bmatrix} \varepsilon \\ 0 \end{bmatrix}.$$

于是

$$\left\| \frac{\mathbf{y}_k}{\|\mathbf{y}_k\|_\infty} - \mathbf{v} \right\|_\infty = |\varepsilon|,$$

即归一化后的结果在无穷范数意义下与 $\mathbf{v}$（从而也与 $\mathbf{v}$ 的某个倍数）相差不超过 $|\varepsilon|$。

Solution 3.5. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-3-inverse-iteration.ipynb`.

Solution 3.6. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-3-inverse-iteration.ipynb`.

4 Krylov subspaces

Solution 4.1. 令

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

(a) 先取种子向量 $\mathbf{u} = \mathbf{e}_1 = [1,0,0,0]^T$。因为 $\mathbf{A}$ 的作用就是循环置换，所以

$$\mathbf{A}\mathbf{u} = \mathbf{A}\mathbf{e}_1 = \mathbf{e}_4, \qquad \mathbf{A}^2\mathbf{u} = \mathbf{A}\mathbf{e}_4 = \mathbf{e}_3.$$

于是

$$K_3 = \big[\mathbf{u}, \mathbf{A}\mathbf{u}, \mathbf{A}^2\mathbf{u}\big] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

(b) 再取 $\mathbf{u} = \mathbf{1} = [1,1,1,1]^T$。这里有个小观察：$\mathbf{A}\mathbf{1} = \mathbf{1}$，所以 $\mathbf{A}^2\mathbf{1} = \mathbf{1}$。因此

$$K_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

所以 $K_3 = \operatorname{span}\{\mathbf{1}\}$，也就是秩为 1。

Solution 4.2. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-4-krylov-subspaces.ipynb`.

Solution 4.3. 由定义

$$K_m = \operatorname{span}\{\, \mathbf{u}, \mathbf{A}\mathbf{u}, \ldots, \mathbf{A}^{m-1}\mathbf{u} \,\},$$

若 $\mathbf{x} \in K_m$，则存在系数 $c_0, \ldots, c_{m-1}$ 使得

$$\mathbf{x} = \sum_{j=0}^{m-1} c_j \mathbf{A}^j \mathbf{u} = p(\mathbf{A})\mathbf{u},$$

其中 $p(z) = \sum_{j=0}^{m-1} c_j z^j$ 是次数不超过 $m-1$ 的多项式。

Solution 4.4. 对 Function 8.4.7（Arnoldi）来说，第 j 次外循环的主要开销为：

1) 计算一次稀疏矩阵-向量乘法 $\mathbf{v} = \mathbf{A}\mathbf{q}_j$，代价约为 $c\,n$ flops；

2) 正交化（内循环 $i = 1, \ldots, j$）每一步包含一次内积与一次向量更新，均为 $O(n)$；若按每步约 $4n$ flops 估计，则总计约 $4nj$ flops；

3) 再做一次范数 $\|\mathbf{v}\|$ 与一次归一化更新，均为 $O(n)$（合并记作 $O(n)$）。

因此第 j 步代价为 $cn + O(nj)$，累加 $j = 1, \ldots, m$ 得

$$\text{flops}(m) \;=\; cn\,m + 4n \sum_{j=1}^{m} j + O(nm) \;=\; cn\,m + 2nm(m+1) + O(nm),$$

其渐近主导项为 $O(nm^2)$（当 m 增大时）以及 $O(cnm)$（来自稀疏 matvec）。

Solution 4.5. (a) 编程题,代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-4-krylov-subspaces.ipynb`.

(b) 取 $\mathbf{u} = [1, 1, 1, 1]^T$ 时会出现 breakdown，直观上是因为 Krylov 子空间退化成一维：$\mathbf{Au}$ 与 $\mathbf{u}$ 共线，正交化后剩余向量会变成零向量，接下来归一化时需要除以 0，从而产生 NaN，算法就无法继续。

Solution 4.6. 记 $\mathbf{v}_0 = \mathbf{A}\mathbf{q}_j$。按 Arnoldi 的定义，

$$H_{ij} = \mathbf{q}_i^* \mathbf{v}_{i-1}, \qquad \mathbf{v}_{i-1} = \mathbf{v}_0 - \sum_{k=1}^{i-1} H_{kj} \mathbf{q}_k.$$

而 Function 8.4.7 采用递推

$$S_{ij} = \mathbf{q}_i^* \Big(\mathbf{v}_0 - \sum_{k=1}^{i-1} S_{kj} \mathbf{q}_k \Big), \qquad i = 1, \ldots, j.$$

用归纳法：当 $i = 1$ 时，$S_{1j} = \mathbf{q}_1^* \mathbf{v}_0 = H_{1j}$. 设对 $k < i$ 已有 $S_{kj} = H_{kj}$，则 $\mathbf{v}_0 - \sum_{k=1}^{i-1} S_{kj} \mathbf{q}_k = \mathbf{v}_0 - \sum_{k=1}^{i-1} H_{kj} \mathbf{q}_k = \mathbf{v}_{i-1}$, 从而 $S_{ij} = \mathbf{q}_i^* \mathbf{v}_{i-1} = H_{ij}$. 故对所有 i 成立 $S_{ij} = H_{ij}$，两种实现数学上等价。

Solution 4.7. (a) Arnoldi 恒等式 (8.4.6) 可以写成

$$\mathbf{A}\mathbf{Q}_m = \mathbf{Q}_{m+1} \mathbf{H}_m,$$

其中 $\mathbf{Q}_{m+1} = [\mathbf{Q}_m, \mathbf{q}_{m+1}]$，$\mathbf{H}_m$ 是 $(m+1) \times m$ 的上 Hessenberg 矩阵。两边左乘 $\mathbf{Q}_m^*$ 得

$$\mathbf{Q}_m^* \mathbf{A} \mathbf{Q}_m = \mathbf{Q}_m^* \mathbf{Q}_{m+1} \mathbf{H}_m = \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \end{bmatrix} \mathbf{H}_m \;=\; \widetilde{\mathbf{H}}_m,$$

这里 $\widetilde{\mathbf{H}}_m$ 就是把 $\mathbf{H}_m$ 的最后一行去掉后得到的 $m \times m$ 矩阵。

(b) 如果把特征值问题限制在 K_m 上，可以取 $\mathbf{x} \approx \mathbf{Q}_m\mathbf{z}$。再从 $\mathbf{A}\mathbf{x} \approx \lambda\mathbf{x}$ 两边左乘 $\mathbf{Q}_m^*$，得到

$$\mathbf{Q}_m^*\mathbf{A}\mathbf{Q}_m\mathbf{z} \approx \lambda\mathbf{z}.$$

再用上一小问得到的 $\mathbf{Q}_m^*\mathbf{A}\mathbf{Q}_m = \widetilde{\mathbf{H}}_m$，就得到低维的近似特征值问题

$$\widetilde{\mathbf{H}}_m\mathbf{z} \approx \lambda\mathbf{z}.$$

(c) 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-4-krylov-subspaces.ipynb`.

5 GMRES

Solution 5.1. 取

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{b} = \mathbf{e}_1.$$

(a) $\mathbf{A}\mathbf{x} = \mathbf{b}$ 由观察得精确解

$$\mathbf{x} = \mathbf{e}_2 = [0, 1, 0, 0]^T.$$

(b) 对 $m = 1, 2, 3$，Krylov 子空间

$$K_m = \operatorname{span}\{\mathbf{b}, \mathbf{A}\mathbf{b}, \ldots, \mathbf{A}^{m-1}\mathbf{b}\} \subset \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_4, \mathbf{e}_3\},$$

所以 $\mathbf{A}\mathbf{x}_m$ 里不可能出现 $\mathbf{e}_1$ 分量来抵消右端 $\mathbf{b} = \mathbf{e}_1$，残差最小化只能选 $\mathbf{x}_m = \mathbf{0}$。于是

$$\mathbf{x}_1 = \mathbf{x}_2 = \mathbf{x}_3 = \mathbf{0}.$$

当 $m = 4$ 时，$K_4 = \mathbb{R}^4$，精确算术下 GMRES 一定能拿到精确解：

$$\mathbf{x}_4 = \mathbf{e}_2.$$

Solution 5.2. 由 $x_m \in K_m$ 知存在次数不超过 $m-1$ 的多项式 p 使得 $\mathbf{x}_m = p(\mathbf{A})\mathbf{b}$（见 8.4.3）。因此

$$\mathbf{r}_m = \mathbf{b} - \mathbf{A}\mathbf{x}_m = \big(\mathbf{I} - \mathbf{A}p(\mathbf{A})\big)\mathbf{b} = q(\mathbf{A})\mathbf{b},$$

其中 $q(z) = 1 - zp(z)$，次数不超过 m，且 $q(0) = 1$。

Solution 5.3. 若 $\operatorname{rank}(K_n) = n$，则 $K_n = \mathbb{R}^n$。方程 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 的精确解 $\mathbf{x}$ 属于 K_n。GMRES 在 K_m 上最小化残差范数；当 $m = n$ 时，由于可达到残差为 0 的点，因此在精确算术下 GMRES 至多 n 步收敛到精确解。

Solution 5.4. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-5-gmres.ipynb`.

Solution 5.5. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-5-gmres.ipynb`.

Solution 5.6. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-5-gmres.ipynb`.

6 MINRES and conjugate gradients

Solution 6.1. 目标是保证 $\|\mathbf{r}_m\|_2/\|\mathbf{b}\|_2 \le 10^{-4}$.

(a) 特征值为 $-100, -99, \ldots, -1, 1, 2, \ldots, 100$，用 (8.6.2)。有 $\kappa_+ = 100, \kappa_- = 100$ 所以 $\|\mathbf{r}_m\|_2/\|\mathbf{b}\|_2 \le (99/101)^{\lfloor m/2 \rfloor}$. 算一下得到可取 $m \ge 922$。

(b) 特征值为 $-100, 1, 2, \ldots, 100$，用 (8.6.2)。$\kappa_+ = 100$ 且 $\kappa_- = 1$，所以 $\|\mathbf{r}_m\|_2/\|\mathbf{b}\|_2 \le (9/11)^{\lfloor m/2 \rfloor}$ 从这里可取 $m \ge 92$。

(c) 特征值为 $1, 2, \ldots, 100$（SPD），用 (8.6.4)。$\kappa = 100$，所以 $\|\mathbf{r}_m\|_2/\|\mathbf{b}\|_2 \le 2(9/11)^m$ 从这里可取 $m \ge 50$。

Solution 6.2. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-6-minres-and-conjugate-gradients.ipynb`.

Solution 6.3. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-6-minres-and-conjugate-gradients.ipynb`.

Solution 6.4. 迭代次数经验上与 $\sqrt{\kappa(\mathbf{A})}$ 成正比。若 $\kappa(\mathbf{A}(t)) = O(t^2)$，则 $\sqrt{\kappa(\mathbf{A}(t))} = O(t)$。因此从 $t = 200$ 到 $t = 300$，迭代次数按比例增长，估计为

$$60 \times \frac{300}{200} = 90.$$

Solution 6.5. 给定实对称矩阵 $\mathbf{A}$ 与向量 $\mathbf{b}$，令 $\phi(\mathbf{u}) = \mathbf{u}^T\mathbf{A}\mathbf{u} - 2\mathbf{u}^T\mathbf{b}$.

(a) 若 $\mathbf{A}\mathbf{x} = \mathbf{b}$，则 $\phi(\mathbf{x}+\mathbf{v}) - \phi(\mathbf{x}) = \mathbf{v}^T\mathbf{A}\mathbf{v}$.

(b) 假设 $\mathbf{x}$ 是一个局部极小点，若 $\mathbf{A}$ 为 SPD，则对任意 $\mathbf{v} \neq 0$ 有 $\mathbf{v}^T\mathbf{A}\mathbf{v} > 0$，所以 ϕ 在 $\mathbf{x}$ 处取得唯一的全局最小值。

(c) 令 $\mathbf{v} = \mathbf{u} - \mathbf{x}$，则 $\phi(\mathbf{u}) - \phi(\mathbf{x}) = \|\mathbf{u} - \mathbf{x}\|_{\mathbf{A}}^2$ （其中 $\|\mathbf{w}\|_{\mathbf{A}} = \sqrt{\mathbf{w}^T\mathbf{A}\mathbf{w}}$）。

(d) 由上一小问可知，最小化 ϕ 等价于最小化 $\|\mathbf{u} - \mathbf{x}\|_{\mathbf{A}}$。而 C G 的 $\mathbf{x}_m$ 正是在仿射 Krylov 子空间 $\mathbf{x}_0 + K_m$ 上，让 $\|\mathbf{u} - \mathbf{x}\|_{\mathbf{A}}$ 最小的解，所以也等价于让 $\phi(\mathbf{u})$ 在 $\mathbf{x}_0 + K_m$ 上最小。

Solution 6.6. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-6-minres-and-conjugate-gradients.ipynb`.

7 Matrix-free iterations

Solution 7.1. 由 (8.7.3) 有 $\text{blur}(X) = B^k X C^k$。对任意 X, Y 与标量 α, $\text{blur}(X+Y) = \text{blur}(X) + \text{blur}(Y)$, $\text{blur}(\alpha X) = \alpha\,\text{blur}(X)$, 故 blur 满足 (8.7.1) 的线性性条件。

Solution 7.2. (a) 置换变换，线性。

(b) 累积求和（下三角全 1 矩阵），线性。

(c) 加常数项，$f(\mathbf{0}) \neq \mathbf{0}$，所以是非线性。

(d) $f(\alpha x) = |\alpha| f(x)$，一般不满足齐次性，所以是非线性。

Solution 7.3. 对标准基 $\mathbf{e}_j$ 计算 $f(\mathbf{e}_j) = \mathbf{A}\mathbf{e}_j$，得到 $\mathbf{A}$ 的第 j 列；拼接所有列即可重构 $\mathbf{A}$。

Solution 7.4. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-7-matrix-free-iterations.ipynb`.

Solution 7.5. (a) 编程题,代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-7-matrix-free-iterations.ipynb`.

(b) 若 $\mathbf{B}$ 为 SPD，则 $\mathbf{B} = \mathbf{Q}\Lambda\mathbf{Q}^T$，从而 $\mathbf{B}^k = \mathbf{Q}\Lambda^k\mathbf{Q}^T$，所以 $\kappa(\mathbf{B}^k) = \kappa(\mathbf{B})^k$。

(c) $k \to \infty$ 时条件数会指数增长，去模糊就会变得非常病态、噪声也会被明显放大；同时 Krylov 方法收敛也更慢，所以通常要配合正则化或早停。

Solution 7.6. (a) $f_i(x) = \sum_{j=1}^{i} x_j$，因此 $f(x+y) = f(x) + f(y)$，$f(\alpha x) = \alpha f(x)$，所以是线性的。

(b) 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-7-matrix-free-iterations.ipynb`.

(c) 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-7-matrix-free-iterations.ipynb`.

8 Preconditioning

Solution 8.1. 设 $\mathbf{M} = \mathbf{R}^T\mathbf{R}$（$\mathbf{R}$ 可逆），令 $\mathbf{C} = \mathbf{R}^{-T}\mathbf{A}\mathbf{R}^{-1}$。则

$$\mathbf{M}^{-1}\mathbf{A} = (\mathbf{R}^T\mathbf{R})^{-1}\mathbf{A} = \mathbf{R}^{-1}\mathbf{R}^{-T}\mathbf{A} = \mathbf{R}^{-1}\mathbf{C}\mathbf{R},$$

即 $\mathbf{M}^{-1}\mathbf{A}$ 与 $\mathbf{C}$ 相似，故二者特征值相同。

Solution 8.2. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-8-preconditioning.ipynb`.

Solution 8.3. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-8-preconditioning.ipynb`.

Solution 8.4. 编程题，代码见 `numerical-analysis-seminar-datahub/solution/code/ch8/8-8-preconditioning.ipynb`.