Inversion of compact perturbations of identity

Alen Alexanderian*

Abstract

We derive an approximate inversion formula for compact peturbations of identity. Specifically, we consider $(I + \mathcal{K})^{-1}$ where $\mathcal{K}$ is a compact positive selfadjoint operator on an infinite-dimensional separable real Hilbert space. This is motivated by an analogous approach in linear algebra.

1 Introduction

Before discussing the Hilbert space setting, we consider a straightforward derivation in linear algebra that motivates the present note. Let $\mathbf{K} \in \mathbb{R}^{n \times n}$ be a symmetric positive semidefinite matrix. It is easy to derive a formula for $(\mathbf{I} + \mathbf{K})^{-1}$ using the spectral decomposition of $\mathbf{K}$. Consider the spectral decomposition $\mathbf{K} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$, where $\mathbf{\Lambda}$ is a diagonal matrix with the (real, non-negative) eigenvalues $\{\lambda_i\}_{i=1}^n$ of $\mathbf{K}$ on its diagonal, and $\mathbf{V}$ is an orthogonal matrix with the eigenvectors $\{\boldsymbol{v}_i\}_{i=1}^n$ as its columns. The eigenvalues are assumed to be in descending order. Clearly, $(\mathbf{I} + \mathbf{K})^{-1} = \mathbf{V}(\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{V}^\top$. This, however, requires knowledge of the all eigenpairs of $\mathbf{K}$. If $\mathbf{K}$ is high-dimensional, this information can be computationally challenging to obtain. In some applications $\mathbf{K}$ has low (numerical) rank and admits a low-rank representation $\mathbf{K} \approx \sum_{j=1}^r \lambda_j \boldsymbol{v}_j \boldsymbol{v}_j^\top$. In this case, it is desirable to derive an approximation to $(\mathbf{I} + \mathbf{K})^{-1}$ that exploits this low-rank structure. This is facilitated by the following relation:

$$(\mathbf{I} + \mathbf{K})^{-1} = \mathbf{V}(\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{V}^\top = \mathbf{I} - \mathbf{V}\mathbf{D}\mathbf{V}^\top. \tag{1.1}$$

where $\mathbf{D} = \mathrm{diag}(\lambda_1/(1+\lambda_1), \cdots, \lambda_n/(1+\lambda_n))$. This is seen by noting $1/(1+\lambda_j) = 1 - \lambda_j/(1+\lambda_j)$, $j = 1, \ldots, n$, from which we obtain $(\mathbf{\Lambda} + \mathbf{I})^{-1} = \mathbf{I} - \mathbf{D}$. The latter implies the second equality in (1.1). Since we assumed that λ_j's are small for $j > r$, we may approximate

$$(\mathbf{I} + \mathbf{K})^{-1} = \mathbf{I} - \mathbf{V}\mathbf{D}\mathbf{V}^\top = \mathbf{I} - \sum_{j=1}^n \frac{\lambda_j}{1+\lambda_j} \boldsymbol{v}_j \boldsymbol{v}_j^\top \approx \mathbf{I} - \sum_{j=1}^r \frac{\lambda_j}{1+\lambda_j} \boldsymbol{v}_j \boldsymbol{v}_j^\top. \tag{1.2}$$

It is worth noting that the derivation leading to the above approximation is typically done using the Sherman–Morrison–Woodbury formula [2, p. 124]. Note also that the error in approximating $(\mathbf{I} + \mathbf{K})^{-1}$ can be quantified easily, and is given by $\| \sum_{j=r+1}^n \frac{\lambda_j}{1+\lambda_j} \boldsymbol{v}_j \boldsymbol{v}_j^\top \|_2 = \frac{\lambda_{r+1}}{1+\lambda_{r+1}}$. The latter also provides a criterion for choosing the truncation level r. Here $\| \cdot \|_2$ denotes the matrix 2-norm.

In this brief note, we seek to derive an analogue of (1.2) in infinite-dimensional Hilbert spaces.

2 Inversion formula for compact perturbation of identity

Let $\mathscr{H}$ be an infinite-dimensional separable real Hilbert space, endowed with inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\| \cdot \| = \langle \cdot, \cdot \rangle^{1/2}$. Assume $\mathcal{K} : \mathscr{H} \to \mathscr{H}$ is a compact selfadjoint positive linear operator. Herein, we rely on the spectral theory for such operators; see, e.g., [3, 1]. Let $\{v_j\}_{j=1}^\infty$ be a complete orthonormal set of eigenvectors of $\mathcal{K}$ and let $\{\lambda_j\}_{j=1}^\infty$ be the corresponding (real non-negative) eigenvalues.

*North Carolina State University, Raleigh, North Carolina, USA. E-mail: alexanderian@ncsu.edu
Last revised: June 20, 2022

Let $z \in \mathscr{H}$ and let $y = (I + \mathcal{K})^{-1}z$. Note that

$$y = \sum_{j=1}^{\infty} \langle y, v_j \rangle v_j = \sum_{j=1}^{\infty} \left\langle (I + \mathcal{K})^{-1}z, v_j \right\rangle v_j = \sum_{j=1}^{\infty} \left\langle z, (I + \mathcal{K})^{-1}v_j \right\rangle v_j = \sum_{j=1}^{\infty} (1 + \lambda_j)^{-1} \langle z, v_j \rangle v_j.$$

That is, $(I + \mathcal{K})^{-1}z = \sum_{j=1}^{\infty} (1 + \lambda_j)^{-1} \langle z, v_j \rangle v_j$. The following result facilitates computing approximations to $(I + \mathcal{K})^{-1}$.

Proposition 1. *We have*

$$(I + \mathcal{K})^{-1}z = z - \sum_{j=1}^{\infty} \frac{\lambda_j}{1 + \lambda_j} \langle z, v_i \rangle v_i, \quad \textit{for all } z \in \mathscr{H}. \tag{2.1}$$

Proof. Let $z \in \mathscr{H}$ be fixed but arbitrary. We have

$$\begin{aligned} (I + \mathcal{K})^{-1}z = \sum_{j=1}^{\infty} (1 + \lambda_j)^{-1} \langle z, v_j \rangle v_j &= \sum_{j=1}^{\infty} \left[1 - \frac{\lambda_j}{1 + \lambda_j} \right] \langle z, v_j \rangle v_j \\ &= \sum_{j=1}^{\infty} \langle z, v_j \rangle v_j - \sum_{j=1}^{\infty} \frac{\lambda_j}{1 + \lambda_j} \langle z, v_j \rangle v_j = z - \sum_{j=1}^{\infty} \frac{\lambda_j}{1 + \lambda_j} \langle z, v_j \rangle v_j. \quad \square \end{aligned}$$

Note that we can write (2.1) in the following form

$$(I + \mathcal{K})^{-1} = I - \sum_{j=1}^{\infty} \frac{\lambda_j}{1 + \lambda_j} v_j \otimes v_j,$$

where $\otimes$ denotes tensor product.[1] Since $\mathcal{K}$ is compact, it follows that $\lambda_j \to 0$, as $j \to \infty$. If the spectral decay is rapid and $\lambda_j \approx 0$ for $j > r$, for some r, then we can approximate

$$(I + \mathcal{K})^{-1} \approx I - \sum_{j=1}^{r} \frac{\lambda_j}{1 + \lambda_j} v_i \otimes v_i. \tag{2.2}$$

Let $\mathcal{S} = (I + \mathcal{K})^{-1}$ and let $\widehat{\mathcal{S}}$ be its approximation from (2.2). If $\mathcal{K}$ has finite rank r, then $\mathcal{S} = \widehat{\mathcal{S}}$. Otherwise, it is straightforward to quantify the approximation error as indicated by the following standard result. We provide a direct proof for completeness.

Proposition 2. *Let $|||\cdot|||$ denote the operator norm induced by the norm $\|\cdot\|$. We have $|||\mathcal{S} - \widehat{\mathcal{S}}||| = \lambda_{r+1}/(1 + \lambda_{r+1})$.*

Proof. Note that for each $z \in \mathscr{H}$,

$$\|\mathcal{S}z - \widehat{\mathcal{S}}z\|^2 = \| \sum_{j=r+1}^{\infty} \frac{\lambda_j}{1 + \lambda_j} \langle z, v_j \rangle v_j \|^2 = \sum_{j=r+1}^{\infty} \left(\frac{\lambda_j}{1 + \lambda_j} \right)^2 \langle z, v_j \rangle^2 \leq \left(\frac{\lambda_{r+1}}{1 + \lambda_{r+1}} \right)^2 \|z\|^2.$$

Thus,

$$\frac{\|(\mathcal{S} - \widehat{\mathcal{S}})z\|}{\|z\|} \leq \frac{\lambda_{r+1}}{1 + \lambda_{r+1}}, \quad \text{for all } z \in \mathscr{H} \setminus \{0\}.$$

Finally, note that

$$\frac{\|(\mathcal{S} - \widehat{\mathcal{S}})v_{r+1}\|}{\|v_{r+1}\|} = \frac{\lambda_{r+1}}{1 + \lambda_{r+1}},$$

which completes the proof. $\square$

References

[1] J. B. Conway. *A course in functional analysis*, volume 96. Springer, 2019.

[2] C. D. Meyer. *Matrix analysis and applied linear algebra*, volume 71. SIAM, 2000.

[3] A. W. Naylor and G. Sell. *Linear operator theory in engineering and science*. Springer, 1982.

[1] For $u, v \in \mathscr{H}$, $u \otimes v$ is the operator $(u \otimes v)z = \langle z, v \rangle u$.