(/course/cs450-s19/flow-session/507256/1/)

1 2 (/course/cs450-s19/flow-session/507256/1/)

Gradient of the Rayleigh Quotient

10 points

For a symmetric matrix A and a nonzero vector $\boldsymbol{x}$, the Rayleigh Quotient is defined by

$$\frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}.$$

In the case where $\boldsymbol{x}_i$ is an eigenvector, then the Rayleigh Quotient will approximate λ_i, the eigenvalue corresponding to that eigenvector.

1. Show that $\nabla \boldsymbol{x}^T \boldsymbol{x} = 2\boldsymbol{x}^T$.
2. Show that $\nabla \boldsymbol{x}^T A \boldsymbol{x} = \boldsymbol{x}^T(A + A^T)$.
3. Using part 1 and part 2, show that

$$\nabla \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} = \frac{2\boldsymbol{x}^T}{||\boldsymbol{x}||_2^2} \left(A - I \frac{\boldsymbol{x}^T A \boldsymbol{x}}{||\boldsymbol{x}||_2^2} \right).$$

4. Let $\boldsymbol{v}_i$ denote the ith eigenvector of A and λ_i denote its corresponding eigenvalue. Show that

$$\nabla \frac{\boldsymbol{x}^T A \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} \bigg|_{\boldsymbol{x}=\boldsymbol{v}_i} = \boldsymbol{0}.$$

Be sure to show all your work and provide justifications for every step to receive full credit.

Please submit your response to this written problem as a PDF file below. You may do either of the following:

- write your response out by hand, scan it, and upload it as a PDF.

  We will not accept unprocessed pictures taken with your phone.

  If you decide to use your phone for scanning, make sure to use an app such as CamScanner (https://www.camscanner.com/) to get a readable PDF. Alternatively, there's a fast and convenient scanner in the Engineering IT office in 2302 Siebel that can just email you a PDF. (It's the Fax-machine-looking thing--not the scanner that's attached to one of the computers.)

- create the PDF using software.

  If you're looking for an easy-ish way to type math, check out TeXmacs (http://texmacs.org/) or LyX (http://www.lyx.org/). Both are installed in the virtual machine. (Under "Applications / Accessories / GNU TeXmacs editor" and "Applications / Office / LyX document processor" respectively.)

Submit your response to each problems in this homework as a separate PDF. If you have multiple PDFs that you need to merge into one, try PDF Split and Merge (http://www.pdfsam.org/download/).

NOTE: Please make sure your solutions are legible and easy to follow. If they are not, we may deduct up to five points *per problem*.

Review uploaded file (blob:https://relate.cs.illinois.edu/a687a266-5be1-415e-9073-782c461ae7cc) · Embed viewer

Uploaded file*

选择文件 未选择任何文件

Your answer is mostly correct. (80.0 %)

The following feedback was provided:

Part 4 is partially correct.

1. We will first look at $\boldsymbol{x}^T\boldsymbol{x}$ component-wise.

$$\boldsymbol{x}^T\boldsymbol{x} = \sum_i^n x_i^2$$

Therefore,

$$\nabla_{x_i} \sum_i^n x_i^2 = 2x_i$$

and $\nabla \boldsymbol{x}^T\boldsymbol{x} = \boldsymbol{x}^T$

2. Similarly, we will consider the component-wise form of $\boldsymbol{x}^T A\boldsymbol{x}$, $\sum_i \sum_j x_i a_{ij} x_j$. Taking the derivative of this expression,

$$\frac{\partial}{\partial x_i} \sum_i \sum_j x_i a_{ij} x_j + \frac{\partial}{\partial x_j} \sum_i \sum_j x_i a_{ij} x_j$$
$$= \sum_j a_{ij} x_j + \sum_i x_i a_{ij}$$
$$\sum_j x_j a_{ji} + \sum_i x_i a_{ij}$$

Therefore, $\nabla \boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{x}^T(A + A^T)$

3. If we let $f(\boldsymbol{x}) = \boldsymbol{x}^T A\boldsymbol{x}$ and $g(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{x}$, then $\nabla \frac{\boldsymbol{x}^T A\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}} = \nabla \frac{f}{g}$.

   Using quotient rule, $\nabla \frac{f}{g} = \frac{f'g - g'f}{g^2}$. From part 1 and part 2, we have f' and g'.

$$
\begin{aligned}
\nabla \frac{f}{g} &= \frac{\boldsymbol{x}^T(A+A^T)\cdot \boldsymbol{x}^T\boldsymbol{x} - \boldsymbol{x}^T A\boldsymbol{x}\cdot 2\boldsymbol{x}^T}{||\boldsymbol{x}||_2^4} \\
&= \frac{2\boldsymbol{x}^T A\cdot \boldsymbol{x}^T\boldsymbol{x} - (2\boldsymbol{x}^T I)\boldsymbol{x}^T A\boldsymbol{x}}{||\boldsymbol{x}||_2^4} \\
&= 2\boldsymbol{x}^T\frac{A\cdot \boldsymbol{x}^T\boldsymbol{x} - I\boldsymbol{x}^T A\boldsymbol{x}}{||\boldsymbol{x}||_2^4} \\
&= \frac{2\boldsymbol{x}^T}{||\boldsymbol{x}||_2^2}\left(\frac{A||\boldsymbol{x}||_2^2 - I\boldsymbol{x}^T A\boldsymbol{x}}{||\boldsymbol{x}||_2^2}\right) \\
&= \frac{2\boldsymbol{x}^T}{||\boldsymbol{x}||_2^2}\left(A - I\frac{\boldsymbol{x}^T A\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}\right)
\end{aligned}
$$

4. Consider the following simplification of $\b x^TA

$$
\begin{aligned}
\boldsymbol{x}^T A &= ((\boldsymbol{x}^T A)^T)^T \\
&= (A^T\boldsymbol{x})^T \\
&= (A\boldsymbol{x})^T
\end{aligned}
.
$$

Expanding the expression in and substituting v_i and λ_i we find (3)

$$
\begin{aligned}
\frac{2\boldsymbol{x}^T}{||\boldsymbol{x}||_2^2}\left(A - I\frac{\boldsymbol{x}^T A\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}\right) &= \\
\frac{2(\boldsymbol{x}^T A)\boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T(\boldsymbol{x}^T A)\boldsymbol{x}}{||\boldsymbol{x}||_2^4} &= \\
\frac{2\lambda_i\boldsymbol{v}_i^T\boldsymbol{v}_i^T\boldsymbol{v}_i - 2\boldsymbol{v}_i^T\lambda_i\boldsymbol{v}_i^T\boldsymbol{v}_i}{||\boldsymbol{v}_i||_2^4} &= \\
\frac{2\lambda_i\boldsymbol{v}_i^T\boldsymbol{v}_i^T\boldsymbol{v}_i - 2\lambda_i\boldsymbol{v}_i^T\boldsymbol{v}_i^T\boldsymbol{v}_i}{||\boldsymbol{v}_i||_2^4} &= \\
&= \mathbf{0}
\end{aligned}
$$