[Special Topics] Homework 3
Cs 4301.001
Sudarshana Jagadeeshi

1-1
We can rewrite $x^T Ax$ as $x^T tx$ where t is an eigen of A. Then the problem becomes:

min -t
such that t is an eigen of A

$$\mathcal{L} = x^T A x - \lambda(x^T x - 1) = x^T A x + \lambda x^T x - \lambda$$

$$\frac{\partial \mathcal{L}}{\partial x} = 2Ax + 2\lambda x = 0 \qquad (A + \lambda I)x = 0$$

But since $x \neq 0$, $A + \lambda I = 0$

$$x^T(A + \lambda I)x - \lambda = 0$$

$$g(\lambda) = -\lambda$$

The dual is maximizing g(λ) over λ, such that -λ is an eigenvalue of A, which we get from rearranging 2Ax+ 2λx.

1-2
We can't use Slater's because it requires the problem to be convex. The solution to the primal problem is the largest eigenvalue. We know that such an largest eigenvalue exists, because xT * A * x is psd and symmetric, meaning there must exist n distinct eigenvalues.

1-3
We rewrite to get

min g(λ) = λ s.t. Ax+λx=0

$$\mathcal{L}(\lambda, \mu) = \lambda + \mu(Ax + \lambda x) = \lambda + \mu A x + \mu \lambda x$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 1 + \mu x = 0$$

subs. $\Rightarrow \mu A x + \lambda(0) = \mu A x$

$$\min_{\mu} \mu A x \quad \text{such that} \quad \mu x = -1$$

2-1
You can divide this subproblem into two parts- the original objective and the new g_k sum term.

>> The added g_k term is a nonnegative weighted sum of convex functions. The log function is concave, so it's negative is convex.
>> The original objective is convex because the norm squared preserves convexity because it just element squares every term (like a parabola,which we know to be convex). The inside of the norm is affine in lambda.
>> When we add these two parts together, it is another nonnegative weighted sum, so the whole thing is convex.

2-2
https://colab.research.google.com/drive/1SmeaVGSKZXlMN_02Q5jsEJDpqBN05osD?usp=sharing

2-3
If the norm between subsequent iterations is sufficiently small, we can converge. However, this criteria won’t work for flatter functions and will terminate early.

3-1
https://colab.research.google.com/drive/1YstOaknKEBWIJcxELkGWEvA6rqnjGyKQ?usp=sharing

3-2
You should pick it close to a guess of what the vector could be. For example, for the matrix [[0,1], [1,1]], I would start my guess at [1,1]. But if you pick a vector orthogonal of the eigenvector with dominant eigenvalue, the method will compute 0/0 for its update and fail.

3-3
I think as stated in the slides, you should pick the eigenvector orthogonal to dominant eigenvector. Similarly for the third largest eigen you should pick a vector perpendicular to the other two.

3-4
Yes, it is possible if the iterating vector and all the possible eigenvectors of the matrix become close to orthogonal.