Math in project1-Lipschitz Bound

---

1) Definition：

A differentiable funciton f is L-smooth or has a Lipschitz continuous gradient with constant L>0 if

$$[|\nabla f(x) - \nabla f(z)| \leq L|x - z|]$$

for any x and z. The norm is l2 norm.

---

2)Lipschiz for this objecive function

Our objective is:

$$[f(x) = \frac{1}{2}|k * x - y|^2 + \frac{\lambda}{2}|x|^2]$$

In circular convolution, we can a cirbular matrix to represent the convolution, which is

$$[Ax := k * x]$$

$$[f(x) = \frac{1}{2}|Ax - y|^2 + \frac{\lambda}{2}|x|^2]$$

We take the derivative:

$$\nabla f(x) = A^\top(Ax - y) + \lambda x$$

---

which is the comments in the code:

```
grad = k^T*(k*x - y) + lam*x
```

Using definition of Lipschiz to get L value

According to the dfinition , we need to calculate the difference:

$$\nabla f(x) - \nabla f(z) = A^\top(Ax - y) + \lambda x - \left(A^\top(Az - y) + \lambda z\right)$$

$$\nabla f(x) - \nabla f(z) = A^\top A(x - z) + \lambda(x - z) = (A^\top A + \lambda I)(x - z)$$

Take the norm:

$$|\nabla f(x) - \nabla f(z)| = |(A^\top A + \lambda I)(x - z)| \leq |A^\top A + \lambda I|_2 \, |x - z|$$

Therefore, we get:

$$L = |A^\top A + \lambda I|_2$$

Since the $A^\top A + \lambda I$ is semidefinite, thus the l2 norm should be the largest eigen value.

$$L = \lambda_{\max}(A^\top A + \lambda I) = \lambda_{\max}(A^\top A) + \lambda$$

Proof:

assume we have one eigenvector v for $A^\top A$

$$A^\top A v = \mu v$$

Then we have:

$$(A^\top A + \lambda I)v = A^\top A v + \lambda I v = \mu v + \lambda v = (\mu + \lambda)v$$

Thus:

$$\lambda_{\max}(A^\top A + \lambda I) = \max_i (\mu_i + \lambda) = \max_i \mu_i + \lambda = \lambda_{\max}(A^\top A) + \lambda$$

In circular convolution, we have:

$$A^\top A = (F^{-1} \operatorname{diag}(K) F)^\top (F^{-1} \operatorname{diag}(K) F) = F^{-1} \operatorname{diag}(\overline{K}) \operatorname{diag}(K) F = F^{-1} \operatorname{diag}(|K|^2) F$$

Thus, the largest eigenvalue is:

$$\lambda_{\max}(A^\top A) = \max_\omega |K(\omega)|^2$$

$$L = \lambda_{\max}(A^\top A) + \lambda = \max_\omega |K(\omega)|^2 + \lambda$$

which is the code:

```
L = float(np.max(np.abs(K) ** 2).real + lam)
```

Use hessian matrix to understand

L is the upper bound of the largest eigenvalue of hessian matrix!

$$\left\|\frac{\nabla f(\boldsymbol{x})-\nabla f(\boldsymbol{z})}{\boldsymbol{x}-\boldsymbol{z}}\right\| \leq \frac{\|\nabla f(\boldsymbol{x})-\nabla f(\boldsymbol{z})\|}{\|\boldsymbol{x}-\boldsymbol{z}\|} \leq L$$

$$\|\nabla^2 f(x)\| \leq L$$

$$\nabla^2 f(x) = H(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{i,j=1}^{n}$$