June 21, 2022

Proposition 1. *Solve linear equation:*

$$Ax = b, \tag{1}$$

where A and b are both large matrix.

Solution. method of steepest descent We can rewrite this problem by minimizing the quadratic function:

$$J(x) = \frac{1}{2} min_x \|Ax - b\|_2. \tag{2}$$

The method of steepest descent tries to find an update in the direction of the steepest descent of the quadratic function (2):

$$J'(x^{(i)}) = Ax^{(i)} - b = -r^{(i)}, \tag{3}$$

where x^i is the solution for i-th iteration, $r^{(i)}$ is the gradient for i-th iteration.

Assume that we have the step α_i for i-th iteration to minimize the loss function J (2),yield

$$\min_{\alpha_i > 0} J\left(x^{(i)} + \alpha_i r^{(i)}\right). \tag{4}$$

The α_i should be chosen such that:

$$\begin{gathered} J'\left(x^{(i)} + \alpha_i r^{(i)}\right) = 0 \\ \Leftrightarrow \left(b - A\left(x^{(i)} + \alpha_i r^{(i)}\right)\right)^T r^{(i)} = 0 \\ \Leftrightarrow \alpha_i = \frac{r^{(i)} r^{(i)}}{r^{(i)T} A r^{(i)}} \end{gathered} \tag{5}$$

Then we get the following steepest descent method for $i = 1, 2, ...$:

$$\begin{array}{ccc} r^{(i)} & := & b - Ax^{(i)}, \\ \alpha_i & := & \frac{r^{(i)T} r^{(i)}}{r^{(i)T} A r^{(i)}}, \\ x^{(i+1)} & = x^{(i)} & +\alpha_i r^{(i)}, \end{array} \tag{6}$$

and the termination condition is $\|r^{(i)}\| \leq \epsilon$, where ϵ is a positive constant. □

Proposition 2. *Using CauchySchwarz inequality to prove Pythagorean theorem.*

Cauchy inequality:

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\| \tag{7}$$

where $\langle \cdot, \cdot \rangle$ is the inner product.

Pythagorean theorem: *As shown in Fig. 1, the area of the square whose side is the hypotenuse (the side opposite the right angle) is equal to the sum of the areas of the squares on the other two sides.*

$$a^2 + b^2 = c^2 \tag{8}$$

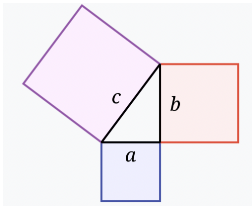

Figure 1: Right triangle

Solution.

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v} \rangle \\
&= \|\mathbf{u}\|^2 + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \|\mathbf{v}\|^2 \quad \text{where } \langle \mathbf{v}, \mathbf{u} \rangle = \overline{\langle \mathbf{u}, \mathbf{v} \rangle} \\
&= \|\mathbf{u}\|^2 + 2\,\mathrm{Re}\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \\
&\le \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v} \rangle| + \|\mathbf{v}\|^2
\end{aligned} \tag{9}$$

Then using Cauchy inequality (7) yields:

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &\le \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2
\end{aligned} \tag{10}$$

From (10), we get the triangle inequality:

$$\|\mathbf{u}+\mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\| \tag{11}$$

The Pythagorean theorem is a special case of triangle inequality, only when the induced angle of $\mathbf{u}$ and $\mathbf{v}$ is right angle. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle = 0$ implies that:

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \tag{12}$$

$\square$