September 8, 2010

Note Title 9/8/2010

A direct method for solving $Ax = b$

Suppose that

$$\{p^{(1)}, p^{(2)}, \ldots, p^{(k)}, \ldots, p^{(n)}\}$$

is a set containing a sequence of n mutually conjugate direction vectors. Then they form a basis for the space $\mathbb{R}^n$.

Hence the solution vector x^* of $Ax = b$ can be written as a linear combination of these basis vectors.

$$x^* = \alpha_1 p^{(1)} + \alpha_2 p^{(2)} + \cdots + \alpha_k p^{(k)} + \cdots + \alpha_n p^{(n)}$$

where the coefficients are given by

$$\alpha_k = \frac{\langle p^{(k)}, b \rangle}{\langle p^{(k)}, p^{(k)} \rangle_A}$$

$A^T = A$

$\therefore \langle x, Ax \rangle = \langle Ax, x \rangle$

$Ax^* = b$ since x^* is a solution.

$$A(\alpha_1 p^{(1)} + \alpha_2 p^{(2)} + \cdots + \alpha_k p^{(k)} + \cdots + \alpha_n p^{(n)}) = b$$

$$\langle p^{(k)}, A(\alpha_1 p^{(1)} + \alpha_2 p^{(2)} + \cdots + \alpha_k p^{(k)} + \cdots + \alpha_n p^{(n)} \rangle = \langle p^{(k)}, b \rangle$$

$$\langle p^{(k)}, \alpha_1 A p^{(1)} \rangle + \langle p^{(k)}, \alpha_2 A p^{(2)} \rangle + \cdots + \langle p^{(k)}, \alpha_n A p^{(k)} \rangle$$

$$+ \cdots + \langle p^{(k)}, \alpha_n A p^{(n)} \rangle = \langle p^{(k)}, b \rangle$$

i.e.

$$\alpha_1 \langle p^{(k)}, A p^{(1)} \rangle + \alpha_2 \langle p^{(k)}, A p^{(2)} \rangle + \cdots + \alpha_k \langle p^{(k)}, A p^{(k)} \rangle$$
$$+ \cdots + \alpha_n \langle p^{(k)}, A p^{(n)} \rangle = \langle p^{(k)}, b \rangle$$

$$\alpha_k \langle p^{(k)}, A p^{(k)} \rangle = \langle p^{(k)}, b \rangle$$

$\alpha_k \langle A p^{(k)}, p^{(k)} \rangle = \langle p^{(k)}, b \rangle$ since A is symmetric

i.e. $\alpha_k \langle p^{(k)}, p^{(k)} \rangle_A = \langle p^{(k)}, b \rangle$ $\boxtimes$

Algorithm

- find the sequence of n conjugate direction vectors $p^{(k)}$
- Compute the coefficients α_k.

The approach is impractical because it would take too much computer time and storage.

We may view the conjugate gradient method as an iterative method.

We carefully choose a small set of conjugate direction vectors $p^{(k)}$ so that we do not need them all to obtain a good approximation to the true solution vector.

algorithm

- Start with an initial guess $x^{(0)}$ to the true solution. (without loss of generality assume $x^{(0)}$ is the zero vector).

- The true solution x^* is also the unique minimizer of

$$f(x) = \frac{1}{2}\langle x, x\rangle_A - \langle x, x\rangle$$

$$= \frac{1}{2}x^T A x - x^T A \quad \text{for } x \in \mathbb{R}^n.$$

So we could take the first basis vector $p^{(1)}$ to be the gradient of f at $x = x^{(0)}$.

i.e. $p^{(1)} = f'(x^{(0)})$

$$= Ax^{(0)} - b$$

$$= -b$$

$$\boxed{\begin{aligned} A^T &= A \\ f'(x) &= Ax - b \end{aligned}}$$

- The other vectors in the basis are now conjugate to the gradient (hence the name of the method)

The k^{th} residual vector

$$r^{(k)} = b - Ax^{(k)}$$

$$Ax^* = b$$

$$0 = b - Ax^*$$

The gradient descent method moves in the direction $r^{(k)}$. Take direction closest to the gradient vector $r^{(k)}$ with direction vector $p^{(k)}$

being conjugate to each other.

$$p^{(k+1)} = r^{(k)} - \frac{\langle p^{(k)}, r^{(k)} \rangle_A}{\langle p^{(k)}, p^{(k)} \rangle_A} p^{(k)}$$

$$x^{(k+1)} = x^{(k)} + \frac{\langle r^{(k)}, r^{(k)} \rangle}{\langle p^{(k)}, p^{(k)} \rangle_A} p^{(k+1)}$$

$Ax = b$

$x^{(0)}$ given

$r^{(0)} = b - Ax^{(0)}$

$p^{(1)} = r^{(0)}$

$$\alpha_1 = \frac{\langle r^{(0)}, r^{(0)} \rangle}{\langle p^{(1)}, p^{(1)} \rangle_A} = \frac{\langle r^{(0)}, r^{(0)} \rangle}{\langle p^{(1)}, A p^{(1)} \rangle}$$

$$x^{(1)} = x^{(0)} + \alpha_1 p^{(1)}$$

$$r^{(1)} = b - A x^{(1)}$$

$$= b - A\left(x^{(0)} + \alpha_1 p^{(1)}\right)$$

$$= \underbrace{b - A x^{(0)}}_{} - \alpha_1 A p^{(1)}$$

$$\boxed{r^{(1)} = r^{(0)} - \alpha_1 A p^{(1)}}$$

$$p^{(2)} = r^{(1)} - \frac{\langle p^{(1)}, r^{(1)} \rangle_A}{\langle p^{(1)}, p^{(1)} \rangle_A} p^{(1)}$$

$$x^{(2)} = x^{(1)} + \frac{\langle r^{(1)}, r^{(1)} \rangle}{\langle p^{(1)}, p^{(1)} \rangle_A} p^{(2)}$$

HW