Iterative Methods in Numerical L.A.

We can be faced with linear algebra problems which are very large and very sparse—certainly too big to handle by standard methods.

Example. The Radiosity Problem in computer graphics. How is "ambient" light distributed in a scene?

Let $B_i =$ light energy leaving triangle $i$. Then

$$B_i = E_i + R_i \sum_{j=1}^{n} B_j F_{ij}$$
where

\[ E_i = \text{energy emitted by triangle } i \]

\[ R_i = \text{reflectivity of patch } i \]

\[ F_{ij} = \text{a "view factor" describing the solid angle subtended by triangle } j \text{ in triangle } i \text{'s sphere of view directions.} \]

\[ F_{ij} = \frac{\cos \Theta_1 \cos \Theta_2}{4\pi s^2} A_1 A_2 \]

where \( A_1, A_2 \) are the areas of the triangles. Note that \( F_{ij} = 0 \) if something is in the way!
In a rendering problem, we know $E_i$ (some of which better be nonzero or there will be no light in the scene!) and the $F_{ij}$ and $R_i$ and are trying to solve for the $B_i$.

We write this as

$$E_i = B_i - R_i \sum_{j=1}^{n} B_j F_{ij}$$

or

$$\dot{E} = (I - [R_1 \ldots R_n][F_{ij}]) \dot{B}$$

In a real scene (for instance, Toy Story 3) is rendered with radiosity throughout, there are a few million triangles, the $F_{ij}$ are very small, and the $R_i$ are (of course) $\ll 1$. 


We can sum up this type of problem:

Solve $Ax = b$.

$A$ is huge, sparse and diagonally dominant.

as a good candidate for an iterative method.

Definition. Given $A, b, x_0$ an iterative method generates a sequence $x_i \rightarrow A^{-1}b$.

Definition. A splitting of $A$ is a decomposition $A = M - K$ with $M$ nonsingular.

Given a splitting, we write

$$Ax = Mx - Kx = b$$

$$\Rightarrow Mx = Kx + b$$

$$\Rightarrow x = M^{-1}Kx + M^{-1}b$$
We then have
\[ R := M^{-1}K, \quad c := M^{-1}b \]
and define our iterative method by
\[ x_{m+1} = Rx_m + c \]
If \( x_{m+1} = x_m \), then we can check that they solve \( Ax = b \) as expected. But when does the sequence converge?

Lemma. Let \( \| \cdot \| \) be an operator norm for matrices. If \( \| R \| < 1 \) then \( x_{m+1} = Rx_m + c \) converges in the corresponding vector norm.

Proof. We know if \( x \) solves \( Ax = b \) then \( x = Rx + c \). So compute
\[
x_{m+1} - x = (Rx_m + c) - (Rx + c) = R(x_m - x)
\]
so
\[
\| x_{m+1} - x \| \leq \| R \| \| x_m - x \| \leq \| R \|^{m+1} \| x_0 - x \| \square.
\]
This convergence criterion doesn't seem very natural (if our preferred operator norm doesn't work, how do we know if another will?). So we now replace it with one that makes more sense.

**Definition.** The spectral radius of $R$ is $\rho(R) = \max |\lambda|$ where $\lambda$ is an eigenvalue of $R$.

**Lemma.** For any operator norm, $\rho(R) \leq \|R\|_\infty$.

Given an $R, \epsilon > 0$ there is an operator norm (depending on $R, \epsilon$) so that

$$\|R\|_\infty \leq \rho(R) + \epsilon.$$ 

**Proof.** Suppose $x_1$ is an eigenvector corresponding to $\lambda$. Then

$$\|R\| = \max_x \frac{\|Rx_1\|}{\|x_1\|} \geq \frac{\|Rx_1\|}{\|x_1\|} \geq |\lambda|.$$
To build a norm, take $S$ so that

$$S^{-1}RS = J$$

is the Jordan form of $R$, or

$$J = \begin{bmatrix}
    \lambda_1 & 1 & 0 & \cdots & 0 \\
    0 & \lambda_2 & 1 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \ddots & \vdots \\
    0 & 0 & \cdots & \lambda_n
\end{bmatrix}$$

we then let

$$D_e = \begin{bmatrix}
    1 & e_2 & \cdots & e_n \\
    0 & 1 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & 1
\end{bmatrix}$$

Now

$$D_e^{-1}J D_e = D_e^{-1} \begin{bmatrix}
    \lambda_1 \lambda_2 0 \cdots 0 \\
    \vdots & \vdots & \ddots & \vdots \\
    \lambda_1 \lambda_2 & \lambda_1 \lambda_2 & \cdots & \lambda_1 \lambda_2
\end{bmatrix}$$

- claim -

$$= \begin{bmatrix}
    \lambda_1 e_{n-1} \lambda_1 e_{n-1} & \cdots & \lambda_1 e_{n-1} \\
    \vdots & \ddots & \vdots \\
    \lambda_1 e_{n-1} & \cdots & \lambda_1 e_{n-1}
\end{bmatrix}$$
We let the vector norm

\[ \|x\|_* = \|SDe^i x\|_\infty \]

Then

\[ \|R\|_* = \max_x \frac{\|Rx\|_*}{\|x\|_*} \]

\[ = \max_x \frac{\|SDe^i R x\|_\infty}{\|SDe^i x\|_\infty} \]

\[ = \max_y \frac{\|SDe^i R (SDe) y\|_\infty}{\|y\|_\infty} \]

\[ = \max_y \|SDe^i R (SDe) y\|_\infty \]

\[ = \max \text{ row sum of the } \epsilon \text{ Jordan form} \]

\[ = \max |\lambda_i| + \epsilon = \rho(R) + \epsilon. \quad \square \]
We can now say

Theorem. The iteration $x_{m+1} = Rx_m + c$ converges to solution of $Ax = b$ for all $x_0$ and all $b$ if $\rho(R) < 1$.

Proof. If $\rho(R) \geq 1$, choose $x_0$ to be an eigenvector with $\lambda \geq 1$. Then

$$(x_{m+1} - x) = R^{m+1}(x_0 - x) = \lambda^{m+1}(x_0 - x)$$

and this can't approach 0. If $\rho(R) < 1$, we can construct an operator norm so $\|R\| < 1$ as above and the iteration converges by Lemma. □

Definition. The rate of convergence of $x_{m+1} = Rx_m + c$ is $r(R) := -\log_{10} \rho(R)$.

This is the number of additional correct digits in solution, per step.
To see this, we compute

\[(X_{m+1} - x) = R (X_m - x)\]

or

\[\|X_{m+1} - x\| \leq \|R\| \|X_m - x\|\]

or

\[\frac{\|X_m - x\|}{\|X_m - x\|} \leq \frac{1}{\|R\|}\]

so

\[\log_{10} \|X_m - x\| - \log_{10} \|X_{m+1} - x\| \geq -\log_{10} \|R\|\]

But if this norm is chosen to be the \(\|\cdot\|_\infty\), so \(\|R\|_\infty \approx p(R) + \varepsilon\), we get

\[\left\{ \log_{10} \|X_m - x\|_\infty - \log_{10} \|X_{m+1} - x\|_\infty \right\} \geq r(R) + O(\varepsilon)\]

- # of correct digits in \(X_m\)
- # of correct digits in \(X_{m+1}\)

Next time: How to choose a splitting.