算法 数学 最优化 矩阵

如何更好的理解共轭梯度方法？

如题

6 个回答 默认排序

日月当空
寻找优雅与有趣

我想从利用最优化方法求解线性方程组的角度来回答，我的这个回答仅限于《数值代数》，至于共轭梯度法更广阔的应用，我还没有学……等下学期学完《最优化方法》再来补充吧。

我第一次接触共轭梯度，是在《数值代数》课本里，共轭梯度法是作为最速下降法的改进方法出现的。一开始也是非常难理解。直到我读完《An Introduction to the Conjugate Gradient Method Without the Agonizing Pain》，然后做了科学计算实践课的大作业后才算搞明白。

我把我大作业里写的这篇报告里的一部分贴出来，希望对大家理解共轭梯度法有帮助。

最速下降法与共轭梯度法

例如：

$$A=\begin{bmatrix}3&2\\2&6\end{bmatrix} b=\begin{bmatrix}2\\-8\end{bmatrix}$$

初始向量：

$$x=[-2,\quad -2]^T$$

求方程$Ax=b$的解

容易看出该方程的解为：$x=[2,\quad -2]^T$

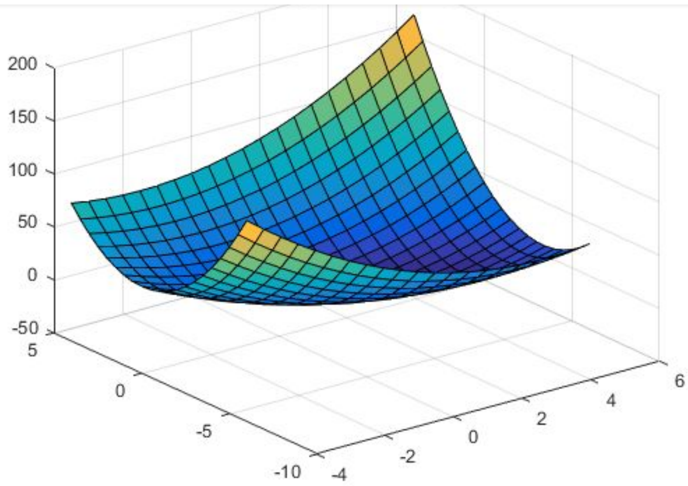
该问题，按照最速下降法和共轭梯度法的观点来看，等价于求下面这个曲面的最小值点：

$$f(x)=\frac{1}{2}x^TAx-b^Tx+c$$

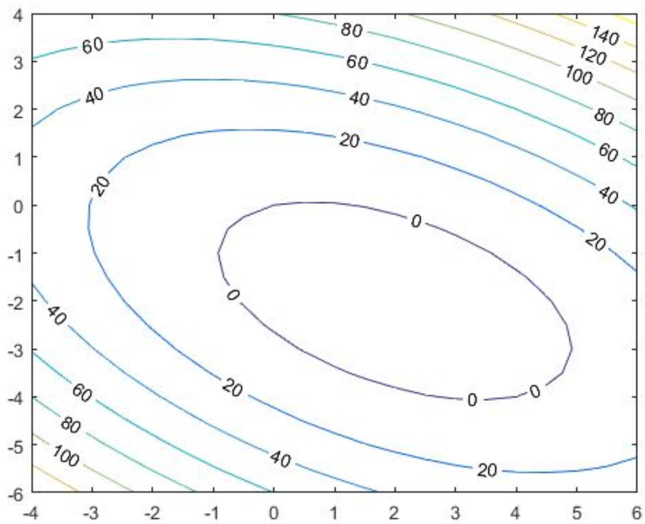
不妨设

$$c=0$$

用MATLAB编程画出曲面如下：

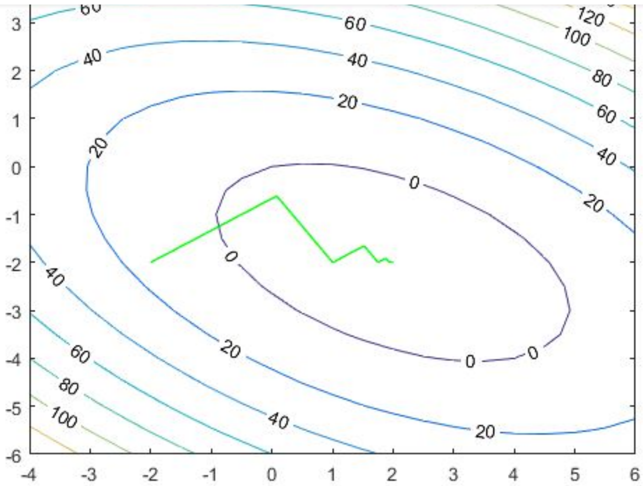


而用等高线图表示是这样的：



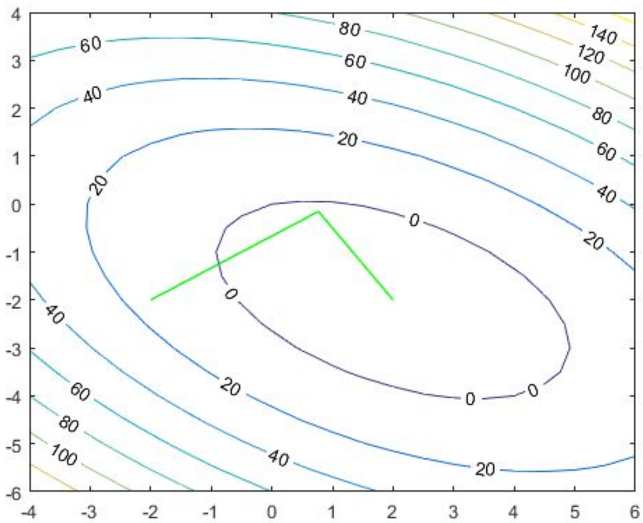
事实上，通过对结果的分析，我发现了共轭梯度法相对于最速下降法，一个在《数值代数》课本里没有提到的好处，我觉得这个好处，实际上才是提出共轭梯度法的重要原因，也是共轭梯度法为什么么要找寻一组A正交基的原因。

事实上，编程输出最速下降法的过程：



可以很自然地看出，最速下降法的每一步都是垂直于上一步，走下降最快的方向，也就是残差方向，而且要在这一步上下降最多，从图中不难发现，在很多中间经过的点上，残差方向往往与前面经过的点残差方向相同，尤其是二维的情况，最为明显，解向量周而复始地在两个互相垂直的方向上前进，但是受到中间点最大步长α的制约，并不能在这个方向上前进很多，往往要再经过一步，才能又回到这一方向上来，继续前进。如果我们，将求解过程中，方向相同的向量加起来，也就是说，将最初一步的绝对误差在n个坐标轴下分解，在求解过程中每一步都走其中一个坐标轴上的分量，求方程组的解，难道不能“n步到位”吗？

也就是说，在二维情况下，是这样的。



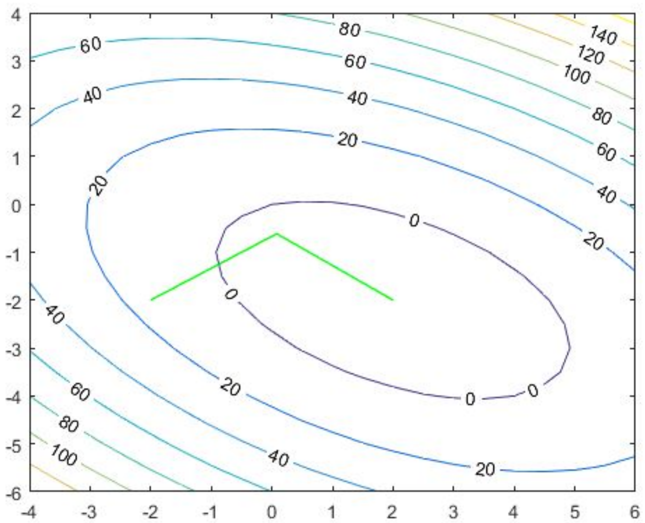
步完成后$d_i^T \perp e_{i+1}$，也就是这一步结束后，绝对误差方向垂直于前进方向，那么对于 n 维问题来说，n 步就一定达到结果，那么根据等式$x_{i+1} = x_i + \alpha_i e_i$，得到$e_i^T(r_i + \alpha_i e_i) = 0$, 然后步长就应取$\alpha_i = \frac{-d_i^T e_i}{d_i^T d_i}$, 但是，这却是不可能完成的，因为我们不可能在不知道的情况下，就得到α_i。

不过，在每一步，我们虽然不知道 e_i，但我们却知道r_i，如果要求d_i A 正交于e_{i+1}的话，那么我们就得到了公式$\alpha_i = \frac{-d_i^T A e_i}{d_i^T A d_i}$，这里有一个很重要的关系$r_i = b - Ax_i = A(x - x_i) = -Ae_i$, 这样，这个困难就被解决了。按照下面的公式就可以计算α_i:

$$\alpha_i = \frac{d_i^T r_i}{d_i^T A d_i}$$

而这就要求，d_i与d_j之间 A 正交，也就是$d_i^T A d_j = 0$我认为，这就是共轭梯度法的想法来源。

而用共轭梯度法演示求解过程，我们可以看到，二维情况下，每一步A正交，两步到位：



在n维情况下，共轭梯度法实现了“n步到位”。

事实上，在我看来，共轭梯度法就是通过仿射变换将整个空间扭曲成A正交的空间，然后实现在直角坐标空间下不可能实现的改进的最速下降方法。

以我个人观点来看，任何方法都不是凭空产生的。由于前人的方法存在不足或者说有值得改进的地方，在对前人方法有充分理解的前提下，后人进行改进，从而提出新方法。但是，随着时间的流逝，新方法的提出者当时所面临的问题有的很难以书面形式保存下来，对问题的观察与思考过程就难了。等到新方法被写进教科书，学生面对的就是干巴巴的结果——方法是什么，而为什么，是怎么提出来的，教材里有的时候没有，这就让学习者头疼。

感谢《An Introduction to the Conjugate Gradient Method Without the Agonizing Pain》，作者在