ACM 100b

Matrix adjoints and linear systems

Dan Meiron

Caltech

January 7, 2014

Solving inhomogeneous systems

- Consider an inhomogeneous linear system of the form

$$\boldsymbol{x}' = A(z)\boldsymbol{x} + \boldsymbol{f}(z), \quad \boldsymbol{x}(z_0) = \boldsymbol{x}_0$$

- In general we cannot solve the homogeneous problem here in contrast with the first order case or with the case of constant coefficient matrices.
- We will assume here that we have determined a solution of the homogeneous system.
- We will also associate with the inhomogeneous problem, a second system of equations called the *adjoint system* defined by

$$\boldsymbol{y}' = -A^T(z)\boldsymbol{y}$$

where A^T is the transpose of $A(z)$.

The adjoint system

- Now suppose we know a solution of the homogeneous adjoint system
- Call that solution $\boldsymbol{y}_1(z)$.
- Then

$$
\begin{aligned}
(\boldsymbol{y}_1^T \boldsymbol{x})' &= \boldsymbol{y}_1^T \boldsymbol{x}' + \boldsymbol{y}_1'^T \boldsymbol{x} \\
&= \boldsymbol{y}_1^T (A\boldsymbol{x} + \boldsymbol{f}) - \boldsymbol{y}_1^T A\boldsymbol{x} \\
&= \boldsymbol{y}_1^T \boldsymbol{f}
\end{aligned}
$$

- This is a first order equation that we can integrate directly to get

$$
\boldsymbol{y}_1^T \boldsymbol{x} = \boldsymbol{y}_1^T(z_0)\boldsymbol{x}_0 + \int_{z_0}^{z} \boldsymbol{y}_1^T(t)\boldsymbol{f}(t)dt.
$$

The adjoint system

- Note that the expression

$$\boldsymbol{y}_1^T \boldsymbol{x} = \boldsymbol{y}_1^T(z_0)\boldsymbol{x}_0 + \int_{z_0}^{z} \boldsymbol{y}_1^T(t)\boldsymbol{f}(t)dt.$$

  can be written in scalar form and this is

$$\sum_{k=1}^{n} y_{1k}x_k = \boldsymbol{y}_1^T(z_0)\boldsymbol{x}_0 + \int_{z_0}^{z} \boldsymbol{y}_1^T(t)\boldsymbol{f}(t)dt,$$

  where we set

$$\boldsymbol{y}_1^T = (y_{11}\ y_{12}\ \dots\ y_{1n}).$$

- This establishes a relationship among components of the adjoint solution and the homogeneous solution x_i.

The adjoint system

- So we have one relation between an adjoint solution and the solution vector $\boldsymbol{x}$:

$$\boldsymbol{y}_1^T \boldsymbol{x} = \boldsymbol{y}_1^T(z_0)\boldsymbol{x}_0 + \int_{z_0}^{z} \boldsymbol{y}_1^T(t)\boldsymbol{f}(t)dt.$$

- We can substitute this into the system to reduce the order of the system by one because if we have all $n-1$ components of $\boldsymbol{x}_k$ we can get the last one from the relation above.
- Now if we can get a second solution of the adjoint equation we can reduce the order by two to $n-2$.
- So if a complete solution of the adjoint system is available then one obtains

$$\sum_{k=1}^{n} y_{ik}x_k = \boldsymbol{y}_i^T(z_0)\boldsymbol{x}_0 + \int_{z_0}^{z} \boldsymbol{y}_i^T(t)\boldsymbol{f}(t)dt, \quad i = 1, 2, ..., n.$$

The adjoint system

- This set of relations

$$\sum_{k=1}^{n} y_{ik} x_k = \mathbf{y}_i^T(z_0)\mathbf{x}_0 + \int_{z_0}^{z} \mathbf{y}_i^T(t)\mathbf{f}(t)dt, \quad i = 1, 2, ..., n.$$

  is now an algebraic system of n equations in n unknowns.
- If we can solve this we can solve the complete system.
- Define the matrix

$$\Phi(z) = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{pmatrix}.$$

- The matrix Φ contains all homogeneous solutions of the adjoint equation.

The adjoint system

- Now using this matrix Φ we can write the system

$$\sum_{k=1}^{n} y_{ik} x_k = \boldsymbol{y}_i^T(z_0)\boldsymbol{x}_0 + \int_{z_0}^{z} \boldsymbol{y}_i^T(t)\boldsymbol{f}(t)dt, \quad i = 1, 2, ..., n.$$

  as

$$\Phi(z)\boldsymbol{x} = \Phi(z_0)\boldsymbol{x}_0 + \int_{z_0}^{z} \Phi(t)\boldsymbol{f}(t)dt.$$

- Since the adjoint solutions we found are linearly independent, the matrix $\Phi(z)$ has an inverse.
- So

$$\boldsymbol{x} = \Phi^{-1}(z)\Phi(z_0)\boldsymbol{x}_0 + \Phi^{-1}(z)\int_{z_0}^{z} \Phi(t)\boldsymbol{f}(t)dt.$$

- This looks very much like the solution we got for first order scalar equations

The adjoint system

- You can verify that Φ solves the adjoint equation in the following way:
$$\Phi^{T'} = -A^T\Phi^T$$
- Or we can write it without transposes by taking the transpose of both sides
$$\Phi' = -\Phi A.$$
- The adjoint approach is useful in that if a solution can be found it can be used to reduce the equation even when there is an inhomogeneous term.
- Note you can also use it to directly solve an inhomogeneous problem.
- If you are lucky to find n linearly independent solutions then there is an alternative approach which is variation of parameters for systems.
- We discuss this after illustrating the adjoint approach for constant coefficient systems.