Massachusetts Institute of Technology

Department of Electrical Engineering and Computer Science

6.243j (Fall 2003): DYNAMICS OF NONLINEAR SYSTEMS

by A. Megretski

Lecture 10: Singular Perturbations and Averaging[1]

This lecture presents results which describe local behavior of parameter-dependent ODE models in cases when dependence on a parameter is not continuous in the usual sense.

10.1 Singularly perturbed ODE

In this section we consider parameter-dependent systems of equations

$$\left\{ \begin{array}{rcl} \dot{x}(t) & = & f(x(t),y(t),t), \\ \epsilon\dot{y} & = & g(x(t),y(t),t), \end{array} \right. \tag{10.1}$$

where $\epsilon \in [0,\epsilon_0]$ is a small positive parameter. When $\epsilon > 0$, (10.1) is an ODE model. For $\epsilon = 0$, (10.1) is a combination of algebraic and differential equations. Models such as (10.1), where y represents a set of less relevant, fast changing parameters, are frequently studied in physics and mechanics. One can say that singular perturbations is the "classical" approach to dealing with uncertainty, complexity, and nonlinearity.

10.1.1 The Tikhonov's Theorem

A typical question asked about the singularly perturbed system (10.1) is whether its solutions with $\epsilon > 0$ converge to the solutions of (10.1) with $\epsilon = 0$ as $\epsilon \to 0$. A sufficient condition for such convergence is that the Jacobian of g with respect to its second argument should be a Hurwitz matrix in the region of interest.

Theorem 10.1 *Let* $x_0 : [t_0,t_1] \mapsto \mathbf{R}^n$, $y_0 : [t_0,t_1] \mapsto \mathbf{R}^m$ *be continuous functions satisfying equations*

$$\dot{x}_0(t) = f(x_0(t),y_0(t),t), \quad 0 = g(x_0(t),y_0(t),t),$$

[1]Version of October 15, 2003

where $f:\ \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R} \mapsto \mathbf{R}^n$ *and* $g:\ \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R} \mapsto \mathbf{R}^m$ *are continuous functions. Assume that* f, g *are continuously differentiable with respect to their first two arguments in a neigborhood of the trajectory* $x_0(t), y_0(t)$*, and that the derivative*

$$A(t) = g_2'(x_0(t), y_0(t), t)$$

is a Hurwitz matrix for all $t \in [t_0, t_1]$*. Then for every* $t_2 \in (t_0, t_1)$ *there exists* $d > 0$ *and* $C > 0$ *such that inequalities* $|x_0(t) - x(t)| \leq C\epsilon$ *for all* $t \in [t_0, t_1]$ *and* $|y_0(t) - y(t)| \leq C\epsilon$ *for all* $t \in [t_2, t_1]$ *for all solutions of (10.1) with* $|x(t_0) - x_0(t_0)| \leq \epsilon$*,* $|y(t_0) - y_0(t_0)| \leq d$*, and* $\epsilon \in (0, d)$*.*

The theorem was originally proven by A. Tikhonov in 1930-s. It expresses a simple principle, which suggests that, for small $\epsilon > 0$, $x = x(t)$ can be considered a constant when predicting the behavior of y. From this viewpoint, for a given $\bar{t} \in (t_0, t_1)$, one can expect that

$$y(\bar{t} + \epsilon\tau) \approx y_1(\tau),$$

where $y_1:\ [0, \infty)$ is the solution of the "fast motion" ODE

$$\dot{y}_1(\tau) = g(x_0(\bar{t}), y_1(\tau)), \quad y_1(0) = y(\bar{t}).$$

Since $y_0(\bar{t})$ is an equilibrium of the ODE, and the standard linearization around this equilibrium yields

$$\dot{\delta}(\tau) \approx A(\bar{t})\delta(\tau)$$

where $\delta(\tau) = y_1(\tau) - y_0(\bar{t})$, one can expect that $y_1(\tau) \to y_0(\bar{t})$ exponentially as $\tau \to \infty$ whenever $A(\bar{t})$ is a Hurwitz matrix and $|y(\bar{t}) - y_0(\bar{t})|$ is small enough. Hence, when $\epsilon > 0$ is small enough, one can expect that $y(t) \approx y_0(t)$.

10.1.2 Proof of Theorem 10.1

First, let us show that the interval $[t_0, t_1]$ can be subdivided into subintervals $\Delta_k = [\tau_{k-1}, \tau_k]$, where $k \in \{1, 2, \ldots, N\}$ and $t_0 = \tau_0 < \tau_1 < \cdots < \tau_N = t_1$ in such a way that for every k there exists a symmetric matrix $P_k = P_k' > 0$ for which

$$P_k A(t) + A(t)' P_k < -I \ \ \forall\ t \in [\tau_{k-1}, \tau_k].$$

Indeed, since $A(t)$ is a Hurwitz matrix for every $t \in [t_0, t_1]$, there exists $P(t) = P(t)' > 0$ such that

$$P(t)A(t) + A(t)'P(t) < -I.$$

Since A depends continuously on t, there exists an open interval $\Delta(t)$ such that $t \in \Delta(t)$ and

$$P(t)A(\tau) + A(\tau)'P(t) < -I \ \ \forall\ \tau \in \Delta(t).$$

Now the open intervals $\Delta(t)$ with $t \in [t_0, t_1]$ cover the whole closed bounded interval $[t_0, t_1]$, and taking a finite number of $\bar{t}_k$, $k = 1, \ldots, m$ such that $[t_0, t_1]$ is completely covered by $\Delta(\bar{t}_k)$ yields the desired partition subdivision of $[t_0, t_1]$.

Second, note that, due to the continuous differentiability of f, g, for every $\mu > 0$ there exist $C, r > 0$ such that

$$|f(x_0(t) + \bar{\delta}_x, y_0(t) + \bar{\delta}_y, t) - f(x_0(t), y_0(t), t)| \le C(|\bar{\delta}_x| + |\bar{\delta}_y|)$$

and

$$|g(x_0(t) + \bar{\delta}_x, y_0(t) + \bar{\delta}_y, t) - A(t)\bar{\delta}_y| \le C|\bar{\delta}_x| + \mu|\bar{\delta}_y|$$

for all $t \in \mathbf{R}$, $\bar{\delta}_x \in \mathbf{R}^n$, $\bar{\delta}_y \in \mathbf{R}^m$ satisfying

$$t \in [t_0, t_1], \quad |\bar{\delta}_x - x_0(t)| \le r, \quad |\bar{\delta}_y - y_0(t)| \le r.$$

For $t \in \Delta_k$ let

$$|\delta_y|_k = (\delta_y' P_k \delta_y)^{1/2}.$$

Then, for

$$\delta_x(t) = x(t) - x_0(t), \quad \delta_y(t) = y(t) - y_0(t),$$

we have

$$|\dot{\delta}_x| \le C_1(|\delta_x| + |\delta_y|_k),$$
$$\epsilon|\dot{\delta}_y|_k \le -q|\delta_y|_k + C_1|\delta_x| + \epsilon C_1 \tag{10.2}$$

as long as δ_x, δ_y are sufficiently small, where C_1, q are positive constants which do not depend on k. Combining these two derivative bounds yields

$$\frac{d}{dt}(|\delta_x| + (\epsilon C_1/q)|\delta_y|) \le C_2|\delta_x| + \epsilon C_2$$

for some constant C_2 independent of k. Hence

$$|\delta_x(\tau_{k-1} + \tau)| \le e^{C_3\tau}(|\delta_x(\tau_{k-1})| + (\epsilon C_1/q)|\delta_y(\tau_{k-1})|) + C_3\epsilon$$

for $\tau \in [0, \tau_k - \tau_{k-1}]$. With the aid of this bound for the growth of $|\delta_x|$, inequality (10.2) yields a bound for $|\delta_y|_k$:

$$|\delta_y(\tau_{k-1} + \tau)| \le \exp(-q\tau/\epsilon)|\delta_y(\tau_{k-1})| + C_4(|\delta_x(\tau_{k-1})| + (\epsilon C_1/q)|\delta_y(\tau_{k-1})|) + C_4\epsilon,$$

which in turn yields the result of Theorem 10.1.

10.2 Averaging

Another case of "potentially discontinuous" dependence on parameters is covered by the following "averaging" result.

Theorem 10.2 *Let $f: \mathbf{R}^n \times \mathbf{R} \times \mathbf{R} \mapsto \mathbf{R}^n$ be a continuous function which is τ-periodic with respect to its second argument t, and continuously differentiable with respect to its first argument. Let $\bar{x}_0 \in \mathbf{R}^n$ be such that $f(\bar{x}_0, t, \epsilon) = 0$ for all t, ϵ. For $\bar{x} \in \mathbf{R}^n$ define*

$$\bar{f}(\bar{x}, \epsilon) = \int_0^\tau f(\bar{x}, t, \epsilon).$$

If $d\bar{f}/dx|_{x=0,\epsilon=0}$ is a Hurwitz matrix, then, for sufficiently small $\epsilon > 0$, the equilibrium $x \equiv 0$ of the system

$$\dot{x}(t) = \epsilon f(x, t, \epsilon) \tag{10.3}$$

is exponentially stable.

Though the parameter dependence in Theorem 10.2 is continuous, the question asked is about the behavior at $t = \infty$, which makes system behavior for $\epsilon = 0$ not a valid indicator of what will occur for $\epsilon > 0$ being sufficiently small. (Indeed, for $\epsilon = 0$ the equilibrium $\bar{x}_0$ is not asymptotically stable.)

To prove Theorem 10.2, consider the function $S: \mathbf{R}^n \times \mathbf{R} \mapsto \mathbf{R}^n$ which maps $x(0), \epsilon$ to $x(\tau) = S(x(0), \epsilon)$, where $x(\cdot)$ is a solution of (10.3). It is sufficient to show that the derivative (Jacobian) $\dot{S}(\bar{x}, \epsilon)$ of $\dot{S}$ with respect to its first argument, evaluated at $\bar{x} = \bar{x}_0$ and $\epsilon > 0$ sufficiently small, is a Schur matrix. Note first that, according to the rules on differentiating with respect to initial conditions, $\dot{S}(\bar{x}_0, \epsilon) = \Delta(\tau, \epsilon)$, where

$$\frac{d\Delta(t,\epsilon)}{dt} = \epsilon \frac{df}{dx}(0, t, \epsilon)\Delta(t, \epsilon), \quad \Delta(0, \epsilon) = I.$$

Consider $\bar{D}(t, \epsilon)$ defined by

$$\frac{d\bar{\Delta}(t,\epsilon)}{dt} = \epsilon \frac{df}{dx}(0, t, 0)\bar{\Delta}(t, \epsilon), \quad \bar{\Delta}(0, \epsilon) = I.$$

Let $\delta(t)$ be the derivative of $\bar{\Delta}(t, \epsilon)$ with respect to ϵ at $\epsilon = 0$. According to the rule for differentiating solutions of ODE with respect to parameters,

$$\delta(t) = \int_0^t \frac{df}{dx}(0, t_1, 0)dt_1.$$

Hence

$$\delta(\tau) = d\bar{f}/dx|_{x=0,\epsilon=0}$$

is by assumption a Hurwitz matrix. On the other hand,

$$\Delta(\tau, \epsilon) - \bar{\Delta}(\tau, \epsilon) = o(\epsilon).$$

Combining this with

$$\bar{\Delta}(\tau, \epsilon) = I + \delta(\tau)\epsilon + o(\epsilon)$$

yields

$$\Delta(\tau, \epsilon) = I + \delta(\tau)\epsilon + o(\epsilon).$$

Since $\delta(\tau)$ is a Hurwitz matrix, this implies that all eigenvalues of $\Delta(\tau, \epsilon)$ have absolute value strictly less than one for all sufficiently small $\epsilon > 0$.