A Mathematical Model of Psychotherapy (or, A Review of Math 2)

Jessica Su

One in four adults suffer from a clinically diagnosable mental illness. Yet only a quarter of these seek therapy, and half of those who seek therapy drop out after the first session. If we understand therapy better, we might be able to improve the therapeutic process and reduce the dropout rate, thus reducing the burden of mental illness around the world.

Allegedly the most important factor in therapeutic success is the relationship between the client and the therapist. We describe this relationship using two variables: the happiness of the client and that of the therapist. With this simplification, we can model this relationship using coupled, piecewise-linear differential equations, then solve them analytically or in Mathematica.

The relationship can be modeled by the equations

$$\frac{dT}{dt} = m_T T + b_T + I_T(C) \tag{1}$$

$$\frac{dC}{dt} = m_C C + b_C + I_C(T) \tag{2}$$

In these equations, T and C are the emotional valence (happiness) of the therapist and client. m_T and m_C are (negative) inertial terms, representing how much the actors respond to their own emotional states (or how difficult it is for their partner to change them). $I_T(C)$ and $I_C(T)$ are "influence functions" reflecting the influence of the client on the therapist and vice versa. b_T and b_C represent their baseline emotional valence - ignoring the influence of the other partner, the system reaches an equilibrium when $mX + b = 0$. Therefore, given m, the term b affects the equilibrium value of the actor's happiness.

The influence functions $I_T(C)$ and $I_C(T)$ can be modeled as piecewise linear. The equations we used were

$$I_T(C) = \begin{cases} 0.5C + 0.5 & C \leq 0 \\ C + 0.5 & 0 < C \leq 1 \\ -0.5C + 2 & C > 1 \end{cases}$$

$$I_C(T) = \begin{cases} 5T - 0.1 & T \leq 0 \\ 0.5T - 0.1 & 0 < T \leq 4 \\ -3T + 13.9 & T > 4 \end{cases}$$

Since these equations are piecewise linear, we can solve them in each region and get the local trajectory at any given point (or most points, at least). In each region, the equations are of the form

$$\begin{aligned} \dot{x} &= ax + by + c \quad (3) \\ \dot{y} &= dy + ey + f \quad (4) \end{aligned}$$

These equations can be solved by elimination to get

$$\begin{aligned} x &= c_1 e^{ut} + c_2 e^{\bar{u}t} + \frac{bf - ce}{ae - bd} \quad (5) \\ y &= c_3 e^{ut} + c_4 e^{\bar{u}t} + \frac{cd - af}{ae - bd} \quad (6) \end{aligned}$$

where $u = \frac{1}{2}[tr + \sqrt{tr^2 - 4det}]$, $\bar{u} = \frac{1}{2}[tr - \sqrt{tr^2 - 4det}]$, $tr = a + e$, and $det = ae - bd$. In the event that $tr^2 - 4det$ is negative, u and $\bar{u}$ are complex, so x and y can be written as a sum of sines and cosines.

Within each region there are three possibilities for the trajectories. When $det > 0$ and $tr^2 - det > 0$, they approach a node, since in this case, u and $\bar{u}$ are negative. When $tr^2 - det < 0$, u and $\bar{u}$ are complex, and the solutions oscillate. When $det < 0$, we have $u > 0$ and $\bar{u} < 0$, so the solutions veer towards a saddle point and evolve along a line with slope c_3/c_1.

Substituting our equations in T and C for the equations in a, b, c, d, e, f, we find that $tr = m_C + m_T$, $det = m_C m_T - I_C I_T$, and $tr^2 - 4det = (m_C - m_T)^2 + 4I_C I_T$. (Here I_C and I_T mean the constant factors describing the influence functions in a certain region.) By doing some math we arrive at four main conclusions.

First of all, most trajectories are spirals, since empirically, the inertial terms are smaller than the influence terms. In practice, this means the therapeutic relationship will usually go through ups and downs and rarely move directly to its equilibrium point. (There are saddle points, though.)

Second, increasing the influence of the other person yields the same sorts of trajectories as responding more weakly to your own previous state. This is because the main determinants (pun not intended) of attractor type are det and $tr^2 - 4det$, and in both of these, increasing/decreasing I has the same effect as decreasing/increasing m.

Third, the less an actor is influenced by his own previous states, the more ups and downs the team will go through before reaching a steady state. This is because the amplitude of the oscillations is determined by $tr = m_C + m_T$, which is larger for smaller values of m_C and m_T.

Finally, the actor most responsive to the other ends up happiest. This can be found by allowing $t \to \infty$ in the final equations for T and C and finding the coordinates of the fixed point.

These conclusions should further our understanding of the therapeutic relationship, thus enabling psychologists to improve their technique.

References

[1] L.S. Liebovitch, P.R. Peluso, M.D. Norman, J. Su, J.M. Gottman. 2011. Mathematical Model of the Dynamics of Psychotherapy. Cogn Neurodyn. http://dx.doi.org/10.1007/s11571-011-9157-x

[2] L.S. Liebovitch, P.R. Peluso, J.M. Gottman, J. Su, M.D. Norman. 2011. A Mathematical Model of the Dynamics of Psychotherapy. Psychother Res. (in press)