George Bernard Dantzig and Uncertainty

John R. Birge
The University of Chicago
Themes

• Much of GBD’s work was motivated by a desire to understand and to deal with uncertainty effectively
• This interest lasted throughout his academic career
• His results in the area form the foundation of multiple branches in both optimization and statistics
Outline

• Beginnings: the “homework” papers
• Middle: Aircraft and the generalized network
• Ends and extensions: Dynamics and sampling
• Conclusions
The “Homework” Papers of GBD’s Thesis

• On the nonexistence of tests of “Student’s” hypothesis having power functions independent of $\sigma$, *Annals of Mathematical Statistics* 11 (1940), 186-192.*


Neyman and Pearson Lemma

• Original (sufficiency) by JN/ESP, *Stat. Res. Memoirs* 1936:

Suppose \( f_1, \ldots, f_{m+1} \) functions,

the class \( S \) of \( S \) s.t. \( \int_S f_i(x)dx = c_i, i = 1, \ldots, m \)

and \( S_0 \) of \( S_0 \) s.t. \( \int_{S_0} f_{m+1}(x)dx \geq \int_S f_{m+1}(x)dx \forall S \in S \)

if \( S \in S, \exists k_1, \ldots, k_m \) s.t. \( f_{m+1}(x) \geq \sum_{i=1}^{m} k_i f_i(x), x \in S, \)

\( f_{m+1}(x) \leq \sum_{i=1}^{m} k_i f_i(x), x \notin S, \)

then \( S \in S_0. \)
GBD and Wald Result

• Necessity: If \( S \subseteq S_0 \), then there exists \( k_1, \ldots, k_m \) satisfying the conditions,

\[
\begin{align*}
    f_{m+1}(x) & \geq \sum_{i=1}^{m} k_i f_i(x), \quad x \in S, \\
    f_{m+1}(x) & \leq \sum_{i=1}^{m} k_i f_i(x), \quad x \notin S.
\end{align*}
\]

• Although not in the paper, GBD viewed the result as an optimization problem.
N-P as Optimization

• Consider the sets $S$ as similar to distributions $P$ in a class $\mathcal{P}$. Finding $S_0$ is like solving for distributions $P^*$ that solve:

$$\max_{P \in \mathcal{P}} \int f_{m+1}(x)P(dx)$$

s.t. $\int f_i(x)P(dx) = c_i, \ i = 1, \ldots, m.$

• This is like a moment problem but how to solve it?
Solution by Generalized Programming

Suppose we have points $x_1, \ldots, x_q$ and wonder if $\{x_1, \ldots, x_m\} = S_0$ and consider an additional point $x'$ not in $S_0$, then the problem is:

$$\max \sum_{j=1}^{m} f_{m+1}(x_j)p_j + f_{m+1}(x')p'$$

s.t. $\sum_{i=1}^{m} f_i(x_j)p_j + f_i(x')p' = c_i, \quad i = 1, \ldots, m.$

Can find $k_1, \ldots, k_m$ s.t. $f_{m+1}(x_j) = \sum_{i=1}^{m} k_i f_i(x_j)$ and including an $x'$ is better if $f_{m+1}(x') > \sum_{i=1}^{m} k_i f_i(x_j)$

This is the reduced cost!
LP for NP

• So,

If $S \in S_0$, then $\exists k_1, \ldots, k_m$ s.t.

$$f_{m+1}(x) \geq \sum_{i=1}^{m} k_i f_i(x), x \in S,$$
$$f_{m+1}(x) \leq \sum_{i=1}^{m} k_i f_i(x), x \notin S.$$ 

• The necessity was proven in GBD’s thesis and LP duality and the simplex method were already there!
LP under Uncertainty

• GBD’s original formulation (two-stage l.p. with recourse):

\[
\max_{x, y(\omega)} c^T x + E_\omega [q^T y(\omega)] \\
\text{s. t. } Ax = b,
\]

\[
Tx + Wy(\omega) = \xi(\omega),
\]

\[
x, y(\omega) \geq 0.
\]

• Motivation: allocate aircraft to flights with random demand to maximize revenue.
Ferguson and GBD

• The allocation of aircraft to routes—an example of linear programming under uncertain demand, *Management Science* 3 (1956), pp. 45-73.

• Decisions: $x_{ij}$ – aircraft i to route j
  $t_{ij}$ – capacity of aircraft i on route j
  $y_{jk}$ – passengers on route j, scenario k
  $q_{jk}$ – expected price on route j, scenario k
  $\xi_{jk}$ – demand on route j, scenario k
  $p_k$ – probability of scenario k

• Observation:
  – Second-stage ($y_{jk}$) is either $\sum t_{ij} x_{ij}$ or $\xi_{jk}$ (simple recourse)
  – Can re-define $y$’s to be incremental revenue as a function of $w_j = \sum t_{ij} x_{ij}$
Generalized Network Model

• Revenue on each route:

  \[ w_1 \quad w_2 \quad w_3 \]

• Overall network:

  Gains \((t_{ij})\)

  Planes \((x)\)

  Loadings \((w)\)
Solution Approach

- **Observation:**
  - A basic solution corresponds to an equal number of arcs and nodes (because of the gains)
  - A cycle exists somewhere
  - Can solve for the multipliers in the cycle parametrically

- **Result:** almost as efficient as transportation simplex method (and 10% greater revenue than previous solution)
GBD Uncertainty’s Next Phase
UC/Berkeley        Stanford
Dynamic Stochastic Programs

\[
\begin{align*}
\min & \quad c_1 x_1 + Q_2 \big( x_1 \big) \\
\text{s.t.} & \quad W_1 x_1 = h_1 \\
& \quad x_1 \geq 0
\end{align*}
\]

\[
Q_t \big( x_{t-1,a(k)} \big) = \sum_{\xi_{t,k} \in \Xi_t} \text{prob} \big( \xi_{t,k} \big) Q_{t,k} \big( x_{t-1,a(k)}, \xi_{t,k} \big)
\]

\[
Q_{t,k} \big( x_{t-1,a(k)}, \xi_{t,k} \big) = \min \quad c_t \big( \xi_{t,k} \big) x_{t,k} + Q_{t+1} \big( x_{t,k} \big) \\
\text{s.t.} & \quad W_t x_{t,k} = h_t \big( \xi_{t,k} \big) - T_{t-1} \big( \xi_{t,k} \big) x_{t-1,a(k)} \\
& \quad x_{t,k} \geq 0
\]

GBD insight (with Gerd Infanger/Peter Glynn): Use Monte Carlo with importance sampling to approximate the value functions \((Q\text{'s})\)
Conclusions

• GBD provided foundation for much of optimization beginning with practical problems and decisions under uncertainty
• Some of the key ideas were already apparent to GBD in 1940
• He continued to make contributions to this field throughout his life