Lecture 9: Entropy Methods for Financial Derivatives

Marco Avellaneda
G63.2936.001

Spring Semester 2009
Table of Contents

1. Review of risk-neutral valuation and model selection
2. One-dimensional models, yield curves
3. Fitting volatility surfaces
4. The principle of Maximum Entropy
5. Weighted Monte Carlo
1. Risk-Neutral Valuation and Model Selection
Risk-neutral valuation

Future states of the economy or market are represented by scenarios described with state variables (prices, yields, credit spreads)

\[ X(t) = [X_1(t), X_2(t), ..., X_n(t)] \quad t \geq 0 \]
Securities produce a stream of state-contingent cash-flows…

\[ F(X(T_1)) \quad F(X(T_2)) \quad F(X(T_3)) \]

Present value of future cash-flows along each scenario:

\[ G(X) = \sum_i \delta(t_i, X) F_i(X(t_i)) \]
Arbitrage Pricing Theory

Consider a market with $M$ reference derivative securities, with discounted cash flows

$$G_1(X), G_2(X), \ldots, G_M(X)$$

trading at (mid-market) prices

$$C_1, C_2, \ldots, C_M$$

If we assume no arbitrage opportunities, there exists a pricing probability measure on the set of future scenarios such that

$$C_j = E^P(G_j(X)), \quad j = 1,2,\ldots, M$$
Risk-neutral valuation

Consider the target derivative security that we wish to price

Present value of future cash-flows along each scenario (as specified by term sheet):

$$G(X) = \sum_i \delta(t_i, X) F_i(X(t_i))$$

Fair Value = $E^P \{ G(X) \}$

$$= E^P \left\{ \sum_i \delta(t_i, X) F_i(X(t_i)) \right\}$$

Fair value = expectation of cash-flows, measured in constant dollars
What goes into the selection of a pricing model?

- Known statistical facts about the market under consideration
  - relevant risk factors
  - model for the dynamics of the underlying stocks, rates, spreads

  Gives rise to a set of scenarios and \textit{a-priori} probabilities for these scenarios, or a stochastic process

- Known prices of cash, forwards and reference derivative securities that trade in the same asset class

  Gives rise to calculation of current risk-premia, to take into account the current prices of derivatives in the same asset class (needed for relative-value pricing)
Example 1: The Forward Rate Curve

a system of consistent forward rates

No arbitrage => a single interest rate for each expiration date
APT => an interest rate "curve"

\[ Z(T) = E^P(\delta(X,T)) \]

Present value of $1 paid in T years

\[ F(T) = \frac{1}{Z(T)} \frac{dZ(T)}{dT} \]

Instantaneous forward rate for loan in period \((T, T+dT)\)

No-arbitrage implies the existence of a discount curve, or forward rate curve (interpolation, splines...)
Forward rate curve consistent with ED Futures and Swaps
# Example #2: Equity Options

## May 20, 2000 Call Series - AOL $56.500

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Issue</th>
<th>Strik Price</th>
<th>Intrinsic Value</th>
<th>bid</th>
<th>Ask</th>
<th>Volume</th>
<th>Open Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOE EH</td>
<td>AOL MAY 20, 2000 $ 40.000 CALL</td>
<td>16.5</td>
<td>16.5</td>
<td>17</td>
<td>26</td>
<td>1159</td>
<td></td>
</tr>
<tr>
<td>AOE EV</td>
<td>AOL MAY 20, 2000 $ 42.500 CALL</td>
<td>14</td>
<td>14.125</td>
<td>14.625</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AOE EI</td>
<td>AOL MAY 20, 2000 $ 45.000 CALL</td>
<td>11.5</td>
<td>12</td>
<td>12.5</td>
<td>21</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>AOE EW</td>
<td>AOL MAY 20, 2000 $ 47.500 CALL</td>
<td>9</td>
<td>9.75</td>
<td>10.125</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>AOO EJ</td>
<td>AOL MAY 20, 2000 $ 50.000 CALL</td>
<td>6.5</td>
<td>7.875</td>
<td>8.25</td>
<td>874</td>
<td>2009</td>
<td></td>
</tr>
<tr>
<td>AOO EK</td>
<td>AOL MAY 20, 2000 $ 55.000 CALL</td>
<td>1.5</td>
<td>4.875</td>
<td>5</td>
<td>498</td>
<td>13987</td>
<td></td>
</tr>
<tr>
<td>AOO EL</td>
<td>AOL MAY 20, 2000 $ 60.000 CALL</td>
<td>0</td>
<td>2.562</td>
<td>2.812</td>
<td>2429</td>
<td>58343</td>
<td></td>
</tr>
<tr>
<td>AOO EM</td>
<td>AOL MAY 20, 2000 $ 65.000 CALL</td>
<td>0</td>
<td>1.375</td>
<td>1.5</td>
<td>2060</td>
<td>48997</td>
<td></td>
</tr>
<tr>
<td>AOO EN</td>
<td>AOL MAY 20, 2000 $ 70.000 CALL</td>
<td>0</td>
<td>0.625</td>
<td>0.75</td>
<td>1470</td>
<td>15796</td>
<td></td>
</tr>
<tr>
<td>AOO EO</td>
<td>AOL MAY 20, 2000 $ 75.000 CALL</td>
<td>0</td>
<td>0.375</td>
<td>0.437</td>
<td>463</td>
<td>14290</td>
<td></td>
</tr>
<tr>
<td>AOO EP</td>
<td>AOL MAY 20, 2000 $ 80.000 CALL</td>
<td>0</td>
<td>0.125</td>
<td>0.25</td>
<td>799</td>
<td>8649</td>
<td></td>
</tr>
<tr>
<td>AOO EQ</td>
<td>AOL MAY 20, 2000 $ 85.000 CALL</td>
<td>0</td>
<td>0.062</td>
<td>0.187</td>
<td>16</td>
<td>6600</td>
<td></td>
</tr>
<tr>
<td>AOO ER</td>
<td>AOL MAY 20, 2000 $ 90.000 CALL</td>
<td>0</td>
<td>0.125</td>
<td>0.25</td>
<td>10</td>
<td>1493</td>
<td></td>
</tr>
<tr>
<td>AOO ES</td>
<td>AOL MAY 20, 2000 $ 95.000 CALL</td>
<td>0</td>
<td>0.062</td>
<td>0.125</td>
<td>0</td>
<td>1744</td>
<td></td>
</tr>
<tr>
<td>AOO ET</td>
<td>AOL MAY 20, 2000 $ 100.000 CALL</td>
<td>0</td>
<td>0.062</td>
<td>0.187</td>
<td>10</td>
<td>596</td>
<td></td>
</tr>
<tr>
<td>AOO EA</td>
<td>AOL MAY 20, 2000 $ 105.000 CALL</td>
<td>0</td>
<td>0.062</td>
<td>0.125</td>
<td>0</td>
<td>182</td>
<td></td>
</tr>
</tbody>
</table>

04/24/00 - 2:11p.m. Eastern. Current Stock Quotes are not delayed
Barrier Option

Stock price = $S(T)$, Strike price = $K$
Payoff = $\max(S(T) - K, 0)$ if $\max(S(t)) < H$

$G(X) = e^{-rT} \max(S(T) - K, 0) * 1_{\max(S(t)) < H}$

Need to define a probability on stock price paths
AOL Jan 2001 Options: Implied volatilities on Dec 20, 2000
Market close

<table>
<thead>
<tr>
<th>Strike</th>
<th>ImpliedVol</th>
<th>VarSwap</th>
</tr>
</thead>
<tbody>
<tr>
<td>31.3</td>
<td>96.6191</td>
<td>87.1215</td>
</tr>
<tr>
<td>32.5</td>
<td>94.5071</td>
<td>87.1215</td>
</tr>
<tr>
<td>33.8</td>
<td>88.4581</td>
<td>87.1215</td>
</tr>
<tr>
<td>35</td>
<td>83.9929</td>
<td>87.1215</td>
</tr>
<tr>
<td>37.5</td>
<td>81.7033</td>
<td>87.1215</td>
</tr>
<tr>
<td>40</td>
<td>82.5468</td>
<td>87.1215</td>
</tr>
<tr>
<td>41.3</td>
<td>81.4319</td>
<td>87.1215</td>
</tr>
<tr>
<td>42.5</td>
<td>80.1212</td>
<td>87.1215</td>
</tr>
<tr>
<td>43.8</td>
<td>78.6667</td>
<td>87.1215</td>
</tr>
<tr>
<td>45</td>
<td>80.7064</td>
<td>87.1215</td>
</tr>
<tr>
<td>46.3</td>
<td>78.8035</td>
<td>87.1215</td>
</tr>
</tbody>
</table>

Pricing probability is not lognormal
The AOL "volatility skews" for several expiration dates
Dupire’s Local Volatility Function

Breeden & Lizenberger 1978
Dupire 1992
Derman & Kani 1994

\[ C(K, T) = C^{2.1} \]

Different differentiable function describing call prices

\[ \sigma^2 \left( S, t \right) = \begin{bmatrix} \frac{\partial C}{\partial T} \\ \frac{K^2}{2} \frac{\partial^2 C}{\partial K^2} \end{bmatrix}_{T = t, K = S} \]

Different interpolations of \( C(K_i, T_j) \) give rise to different local vols
Higher dimensions?
Model Selection Issues

- Different interpolation mechanisms for rate curves/ volatility surfaces give rise to different valuations.

- How do we take into account the historical data in conjunction with the choice of model?

- How do we generate stable and easy-to-implement model generation schemes that can be fitted to the prices of many reference derivatives?

- Few parameters (e.g. Stochastic volatility) allows to calibrate to a few reference instruments; many parameters (local volatility surfaces) lead to ill-posed problems.

- Curse of dimensionality: how can we write and calibrate models with many underlying assets (bespoke CDO tranches, multi-asset equity derivatives)?
2. The Principle of Maximum Entropy
Boltzmann’s counting argument

N boxes, Q balls \((Q \gg N)\)

Configuration: an assignment or mapping of each ball to a box (or “state”)

Counting probability distribution associated with a configuration:

\[
p_i = \frac{\text{number of balls in box } i}{N} \quad i = 1, \ldots, N
\]
How many configurations give rise to a given probability?

\[ p_i = \frac{n_i}{Q}, \quad \sum_{i=1}^{N} n_i = Q, \quad i = 1, \ldots, N. \]

\[ \nu(p_1, \ldots, p_N) = \frac{Q!}{n_1! n_2! \ldots n_N!} \]

Number of configurations consistent with \( p \)

Stirling’s approximation

\[ m! \approx \frac{1}{\sqrt{2\pi}} n^{n+1/2} e^{-n} \]

\[ \nu(p_1, \ldots, p_N) \approx Q \cdot \sum_{i=1}^{N} p_i \ln \left( \frac{1}{p_i} \right) = -Q \left( \sum_{i=1}^{N} p_i \ln p_i \right) \quad Q >> N \]
Most likely probability (under constraints)

No constraints:

$$\sum_{i=1}^{N} p_i \ln \left( \frac{1}{p_i} \right) \leq \ln N \quad \text{with equality iff } \quad p_i = 1/N$$

M linear moment constraints:

$$\sum_{i=1}^{N} g_{ij} p_i = c_j \quad j = 1, \ldots, M$$

$$\max \left\{ \sum_{i=1}^{N} p_i \ln \left( \frac{1}{p_i} \right) \left| \sum_{i=1}^{N} g_{ij} p_i = c_j \right. \quad j = 1, \ldots, M \right\}$$
Dual Method

Solve

$$\min_p \left\{ -\sum_i p_i \ln p_i + \sum_{j=1}^{M} \lambda_j \left( \sum_{i=1}^{N} p_i g_{ij} - c_j \right) + \lambda_0 \left( \sum_{i=1}^{N} p_i - 1 \right) \right\}$$

$$-\ln p_i - 1 + \sum_{j=1}^{M} \lambda_j g_{ij} + \lambda_0 = 0 \quad \therefore$$

$$p_i = \frac{1}{Z(\lambda)} \exp \left( \sum_{j=1}^{M} \lambda_j g_{ij} \right), \quad Z(\lambda) = \sum_{i=1}^{N} \exp \left( \sum_{j=1}^{M} \lambda_j g_{ij} \right)$$
Calibration Problem for Equity Derivatives

Given a group, or collection of stocks, build a stochastic model for the joint evolution of the stocks with the following properties:

• The associated probability measure on market scenarios is risk-neutral: all traded securities are correctly priced by discounting cash-flows

• The associated probability measure is such that stock prices, adjusted for interest and dividends, are martingales (local risk-neutrality)

• The model simulates the joint evolution of ~ 100 stocks

• All options (with reasonable OI), forward prices, on all stocks, must be fitted to the model. Number of constraints ~50 to ~1000 or more

• Efficient calibration, pricing and sensitivity analysis
Example: Basket of 20 Biotechnology Stocks (Components of BBH)

<table>
<thead>
<tr>
<th>Ticker</th>
<th>Price</th>
<th>ATM ImVol</th>
<th>Ticker</th>
<th>Price</th>
<th>ATM ImVol</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABI</td>
<td>17.85</td>
<td>55</td>
<td>GILD</td>
<td>30.05</td>
<td>46</td>
</tr>
<tr>
<td>AFFX</td>
<td>17.19</td>
<td>64</td>
<td>HGSI</td>
<td>16.99</td>
<td>84</td>
</tr>
<tr>
<td>ALKS</td>
<td>5.79</td>
<td>106</td>
<td>ICOS</td>
<td>23.62</td>
<td>64</td>
</tr>
<tr>
<td>AMGN</td>
<td>44.1</td>
<td>40</td>
<td>IDPH</td>
<td>43.31</td>
<td>72</td>
</tr>
<tr>
<td>BGEN</td>
<td>35.36</td>
<td>41</td>
<td>MEDI</td>
<td>27.75</td>
<td>82</td>
</tr>
<tr>
<td>CHIR</td>
<td>32.03</td>
<td>37</td>
<td>MLNM</td>
<td>11.8</td>
<td>92</td>
</tr>
<tr>
<td>CRA</td>
<td>10.2</td>
<td>55</td>
<td>QLTI</td>
<td>9.36</td>
<td>64</td>
</tr>
<tr>
<td>DNA</td>
<td>33.27</td>
<td>53.5</td>
<td>SEPR</td>
<td>6.51</td>
<td>84</td>
</tr>
<tr>
<td>ENZN</td>
<td>22.09</td>
<td>81</td>
<td>SHPGY</td>
<td>25.2</td>
<td>47</td>
</tr>
<tr>
<td>GENZ</td>
<td>21.66</td>
<td>56</td>
<td>BBH</td>
<td>81.5</td>
<td>32</td>
</tr>
</tbody>
</table>
Implied Volatility Skews
Multiple Names, Multiple Expirations

AMGN Exp: Oct 00

BGEN Exp: Oct 00

MEDI Exp: Dec 00

Needed:
- 20-dimensional stochastic process
- fits option data (multiple expired)
- martingale property
Multi-Dimensional Diffusion Model

\[ \frac{dS_i}{S_i} = \sigma_i dZ_i + \mu_i dt \quad \mu_i = r - d_i \quad \text{ensures martingale property} \]

\[ dZ_i = \text{Brownian motion increment} \]

\[ E(dZ_i dZ_j) = \rho_{ij} dt \]

1-Dimensional Problems
Dupire: local volatility as a function of stock price \( \sigma = \sigma(S, t) \)
Hull-White, Heston: more factors to model stochastic volatility
Rubinstein, Derman-Kani: implied "trees"

These methods do not generalize to higher dimensions. They are "rigid" in terms of the modeling assumptions that can be made.
Main Challenges in Multi-Asset Models

• Modeling correlation, or co-movement of many assets

• Correlation may have to match market prices if index options are used as price inputs (time-dependence)

• Fitting single-asset implied volatilities which are time- and strike-dependent

• Large body of literature on 1-D models, but much less is known on intertemporal multi-asset pricing models

Beware of ``magic fixes”, e.g. Copulas
Weighted Monte Carlo

Avellaneda, Buff, Friedman, Grandchamp, Kruk: IJTAF 1999

• Build a discrete-time, multidimensional process for the asset price
• Generate many scenarios for the process by Monte Carlo Simulation
• Fit all price constraints using a Maximum-Entropy algorithm
MC with Non-Uniform Probabilities

Avellaneda, Buff, Friedman, Kruk, Grandchamp: IJTAF, 1999

• SDE is used to sample the path space
  \[ dX = \Sigma \cdot dW + B \cdot dt \]

• SDE represents Bayesian prior, e.g. subjective probability

• Reweighted probabilities reflect prices of traded securities - Arrow-Debreu probabilities
**MC with Non-Uniform Probabilities**

Avellaneda, Buff, Friedman, Kruk, Grandchamp: IJTA, 1999

- SDE is used to sample the path space
  \[ dX = \Sigma \cdot dW + B \cdot dt \]
- SDE represents Bayesian prior, e.g. subjective probability

- Reweighted probabilities reflect prices of traded securities - Arrow-Debreu probabilities
Example 1: Discrete-Time Multidimensional Markov Process

Modeled after a diffusion

\[ S_{n+1}^{(i)} = S_n^{(i)} \cdot \left[ 1 + \sigma_n^{(i)} \left( \sum_{j=1}^{N} \alpha_{ij} \xi_{n,j} \right) \sqrt{\Delta t} + \mu_n^{(i)} \Delta t \right] \]

\[ \xi_{n,j} = \text{i.i.d. normals} \]

- Correlations estimated from econometric analysis
- Vols are ATM implied or estimated from data
- Time-dependence, seasonality effects, can be incorporated
Example 2: Multidimensional Resampling

\[ S_{ni} = \text{historical data matrix} \quad n \leq \nu \quad (\text{sample size}) \]

\[ X_{ni} = \frac{S_{ni} - S_{(n-1)i}}{S_{(n-1)i}} \quad Y_{ni} = \frac{X_{ni}}{\sqrt{\sum_{m=1}^{\nu} (X_{mi} - \bar{X}_i)^2}} \]

Use resampled standardized moves to generate scenarios

\[ S^{(i)}_{n+1} = S^{(i)}_n \cdot \left[ 1 + \sigma_{n}^{(i)} Y_{R(n),i} \sqrt{\Delta t} + \mu_{n}^{(i)} \Delta t \right] \]

\[ R(n) = \text{random number between 1 and } \nu \]

\( R(n) \) can be uniform or have temporal correlation
Two draws from the empirical distribution (12/99-12/00)

Simulation consists of sequence of random draws from standardized empirical distribution
Calibration to Option and Forward Prices

- Evaluate Discounted Payoffs of reference instruments along different paths
  \[ g_{ij} = e^{-rT_j} \max\left(S_{i,T_j}^{a_j} - K_j, 0\right) \]

  \( i = 1, \ldots, N \) (number of simulated paths)

  \( j = 1, \ldots, M \) (number of reference instruments)

  \( C_j = \) midmarket price of \( j^{th} \) reference instrument

- Solve

  \[
  \begin{pmatrix}
  C_1 \\
  \vdots \\
  C_M
  \end{pmatrix}
  =
  \begin{pmatrix}
  g_{11} & g_{12} & \cdots & \cdots & g_{1N} \\
  \vdots & \vdots & \ddots & \cdots & \vdots \\
  g_{M1} & \cdots & \cdots & \cdots & g_{MN}
  \end{pmatrix}
  \begin{pmatrix}
  p_1 \\
  p_2 \\
  \vdots \\
  p_N
  \end{pmatrix}
  \]

- Repricing condition

  \[ C_j = E^P(g_j(S)), \quad j = 1, 2, \ldots, M \]
Maximum-Entropy Algorithm

\[ H(p) = - \sum_{i=1}^{N} p_i \log p_i = -D(p \parallel u) \quad u = \left( \frac{1}{N}, \ldots, \frac{1}{N} \right) \]

Algorithm: solve

\[ \max_p H(p) \quad \text{subject to price constraints} \]
\[ \min_p D(p \parallel u) \quad " \]

Stutzer, 1996; Buchen and Kelly, 1997; Avellaneda, Friedman, Holmes, Samperi, 1997; Avellaneda 1998; Cont and Tankov, 2002, Laurent and Leisen, 2002, Follmer and Schweitzer, 1991; Marco Frittelli MEM
Calibrated Probabilities are Gibbs Measures

Lagrange multiplier approach for solving constrained optimization gives rise to $M$-parameter family of Gibbs-type probabilities

$$p_i = p_i = \frac{1}{Z(\lambda)} \exp \left( \sum_{j=1}^{M} \lambda_j g_{ij} \right), \quad i = 1, 2, \ldots, N$$

$$Z(\lambda) = \sum_{i=1}^{N} \exp \left[ \sum_{j=1}^{M} \lambda_j g_{ij} \right]$$

Boltzmann-Gibbs partition function

Unknown parameters
Calibration Algorithm
How do we find the lambdas?

- Minimize in lambda

\[ W(\lambda) = \log Z(\lambda) - \sum_{j=1}^{M} \lambda_j C_j \]

- \( W \) is a convex function

- The minimum is unique, if it exists

- \( W \) is differentiable in C, lambda with explicit gradient

- Use L-BFGS Quasi-Newton gradient-based optimization routine
Boltzmann-Gibbs formalism

\[
\frac{\partial W(\lambda)}{\partial \lambda_j} = E^{P_\lambda}(G_j(X)) - C_j
\]

Gradient = difference between market px and model px

\[
\frac{\partial^2 W(\lambda)}{\partial \lambda_j \partial \lambda_k} = E^{P_\lambda}(G_j(X)G_k(X)) - C_j C_k = Cov^{P_\lambda}(G_j(X), G_k(X))
\]

Hessian = covariance of cash-flows under pricing measure

Numerical optimization with known gradient & Hessian also possible
Least-Squares Version

\[ \chi^2 = \sum_{j=1}^{M} \left( \sum_{i=1}^{N} g_{ij} p_i - C_j \right)^2 = \sum_{j=1}^{M} \left( E^P(g_j(S)) - C_j \right)^2 \]

\[ \min_p \left\{ -H(p) + \frac{\chi^2}{2\epsilon^2} \right\} \]

Max entropy with least-squares constraint

\[ \min_\lambda \left\{ \ln Z(\lambda) + \sum_{j=1}^{M} \lambda_j C_j + \frac{\epsilon^2}{2} \sum_{j=1}^{M} \lambda_j^2 \right\} \]

 Equivalent to adding quadratic term to objective function
Sensitivity Analysis

\[ h(X) = \text{payoff function of "target security"} \]
\[ E_{P\lambda}^\ast(h(X)) = \text{model value of } " \]

\[
\frac{\partial E_{P\lambda}^\ast(h(X))}{\partial C_j} = \frac{\partial E_{P\lambda}^\ast(h(X))}{\partial \lambda_k} \frac{\partial \lambda_k}{\partial C_j}
\]

\[
= \text{Cov}_{P\lambda}^\ast(h(X), g_k(X)) \cdot \left( \frac{\partial C}{\partial \lambda_k} \right)^{-1}_{kj}
\]

\[
= \text{Cov}_{P\lambda}^\ast(h(X), g_k(X)) \cdot \left( \text{Cov}_{P\lambda}^\ast(g\ast(X), g\ast(X)) \right)^{-1}_{jk}
\]
Price-Sensitivities= Regression Coefficients

Solve LS problem:

\[
\min_{\beta, \alpha} \sum_{i=1}^{\nu} p_i \left( h(X_i) - \alpha - \sum_{j=1}^{M} \beta_j G_j(X_i) \right)^2
\]

Uncorrelated to \( g_j(X) \)

\[
h(X) = \alpha + \sum_{j=1}^{M} \beta_j g_j(X) + \epsilon(X)
\]
Minimal Martingale Measure?


- Boltzmann-Gibbs posterior measure with price constraints is not a local martingale

- Remedy: include additional constraints:

  \[ g(S) = \psi(S_{t_1}, ..., S_{t_N})(S_{t_{N+1}} - S_{t_N}) \]

  \[ \psi(S_{t_1}, ..., S_{t_N}) = \text{polynomial function} \]

  Martingale constraint:

  \[ E^P(g(S)) = 0 \quad \text{for all } \psi \]

- Constrained Max-Entropy problem with martingale constraints: Follmer-Schweitzer MEM under constraints

- In practice, use only low-degree polynomials (deg=0 or deg=1)