Модель кредитного портфеля и стресс-тестирование

Департамент рисков ВТБ

Хованский Н.К.

Селезнев В.В.

Суржко Д.А.

План презентации

1. Введение. Принципы моделей кредитного портфеля.

2. Модель кредитного портфеля (распределение Васичека).

3. Модификация распределения потерь и стресс-тест.

4. Модификация Многофакторной модели распределения потерь.

5. Сравнение результатов калибровки с IRB формулой Базель-2.

6. Аспекты практической имплементации модели.

7. Заключение.

© ВТБ 2013

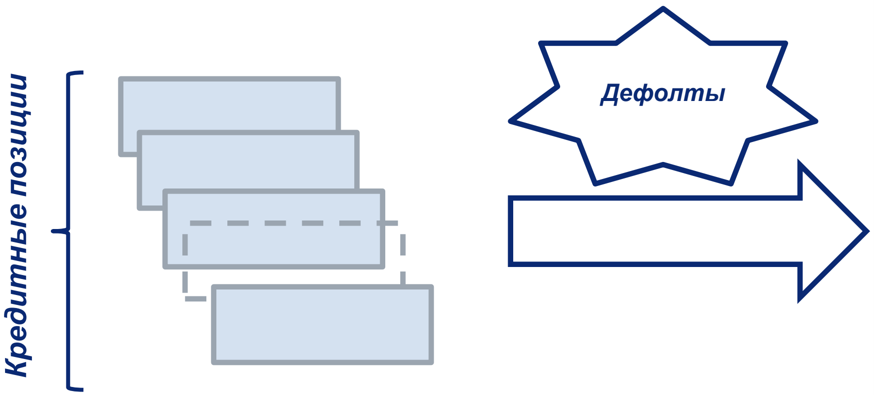
План презентации

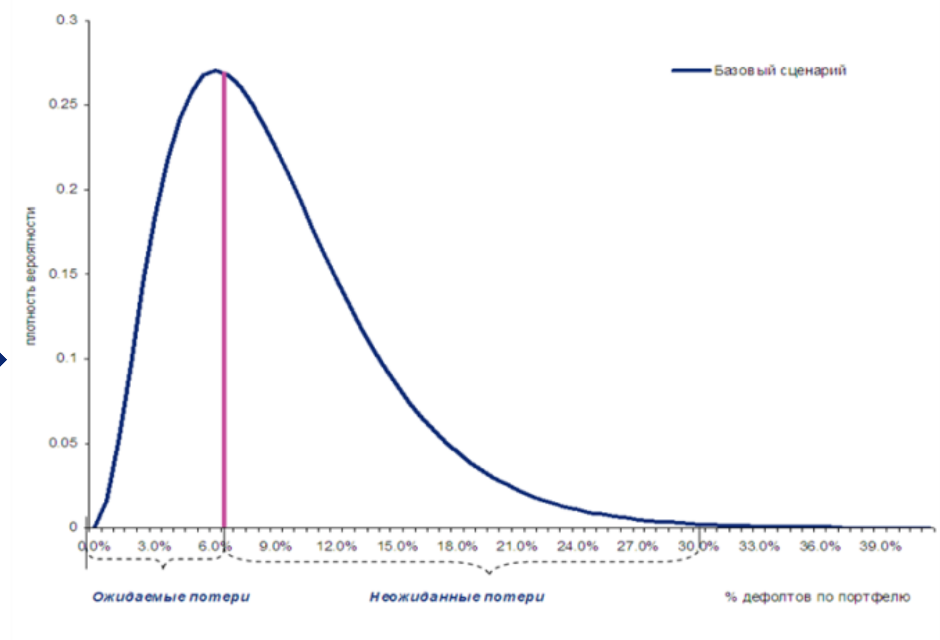
1. **Введение. Принципы моделей кредитного портфеля.**
2. Модель кредитного портфеля (распределение Васичека).
3. Модификация распределения потерь и стресс-тест.
4. Модификация Многофакторной модели распределения потерь.
5. Сравнение результатов калибровки с IRB формулой Базель-2.
6. Аспекты практической имплементации модели.
7. Заключение.

© ВТБ 2013

Введение. Модели кредитного портфеля.

Основная цель Модели Кредитного Портфеля оценка на заданном временном горизонте функции распределения потерь, получаемой вследствие реализации кредитного риска, используя при этом информацию об отдельных заемщиках.





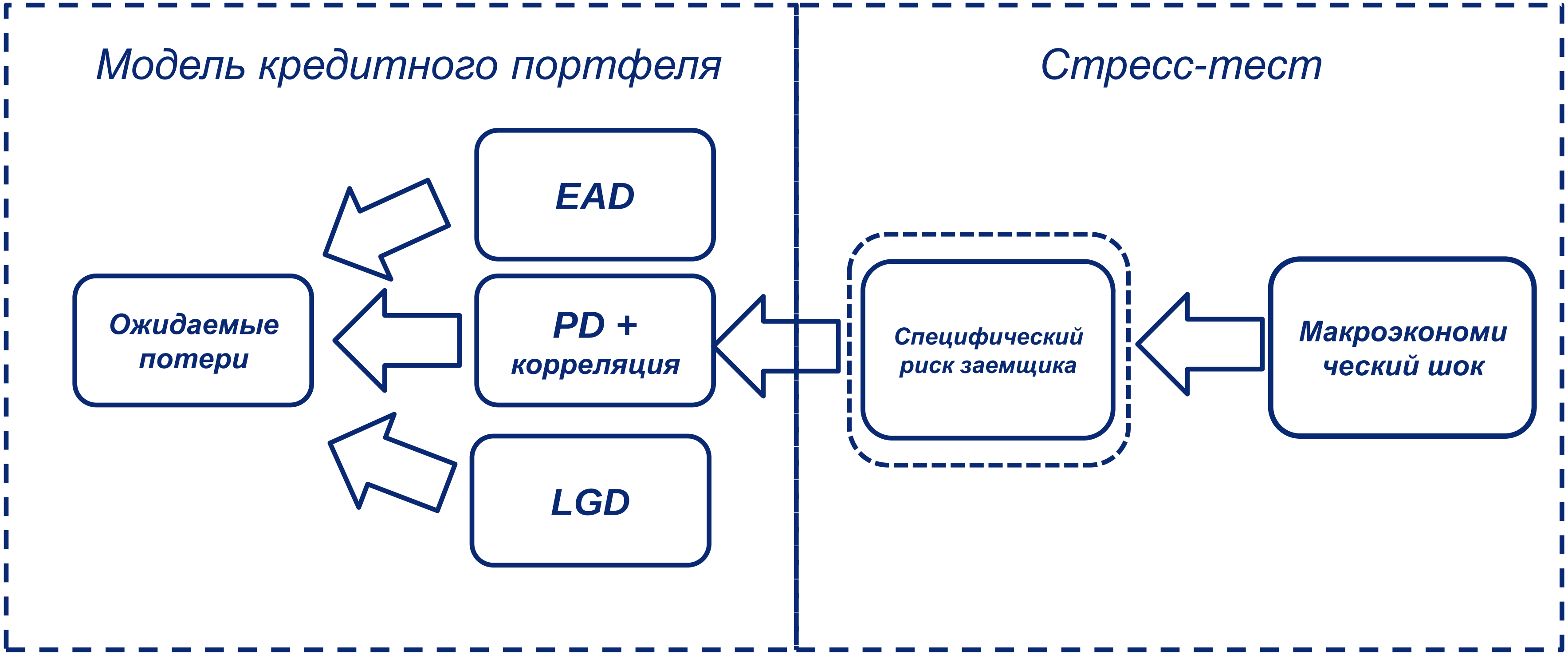
Основном результатом Модели является оценка Ожидаемых и Неожидаемых потерь по портфелю.

- *Ожидаемые потери покрываются за счет резервов, сформированных финансовой организацией.*
- *Неожидаемые потери рассчитываются как разница между 99% квантилем функции распределения потерь и Ожидаемыми потерями, и покрываются за счет экономического капитала кредитной организации.*

© ВТБ 2013

Введение. Модели кредитного портфеля и стресс тест.

Согласно общепринятой практике, функция распределения потерь моделируется на основе набора параметров: PD, LGD, EAD, корреляции между заемщиками.



В дальнейшем предполагается:

1. *Постоянное значение EAD, равное сумме задолженности на момент моделирования.*
2. *Отсутствие корреляции между LGD и PD.*
3. *Постоянное значение LGD.*

© ВТБ 2013

План презентации

1. Введение. Принципы моделей кредитного портфеля.

2. Модель кредитного портфеля (распределение Васичека).

3. Модификация распределения потерь и стресс-тест.

4. Модификация Многофакторной модели распределения потерь.

5. Сравнение результатов калибровки с IRB формулой Базель-2.

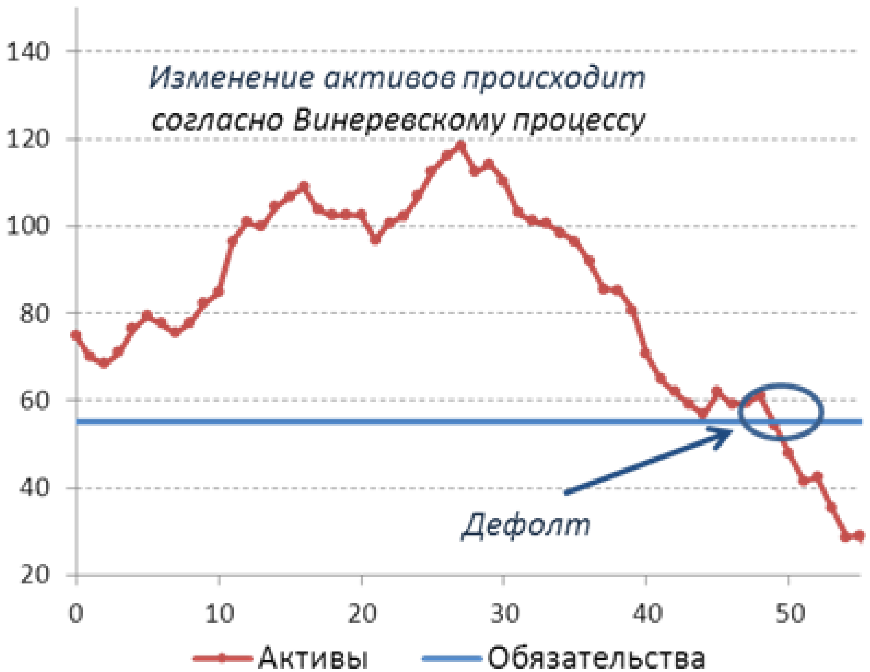
6. Аспекты практической имплементации модели.

7. Заключение.

© ВТБ 2013

Модель кредитного портфеля. (Распределение Васичека.)

Вероятность рассчитывается на основе структурной модели Мертона, и дефолт происходит тогда , когда за заданный промежуток времени активы компании становятся ниже обязательств ($P(A_{iT} < D_i)$)



Изменение активов описывается стохастическим дифференциальным уравнением:

$$dA_i = rA_i dt + \sigma_i A_i dz_i$$

,где z-Винеровский процесс.

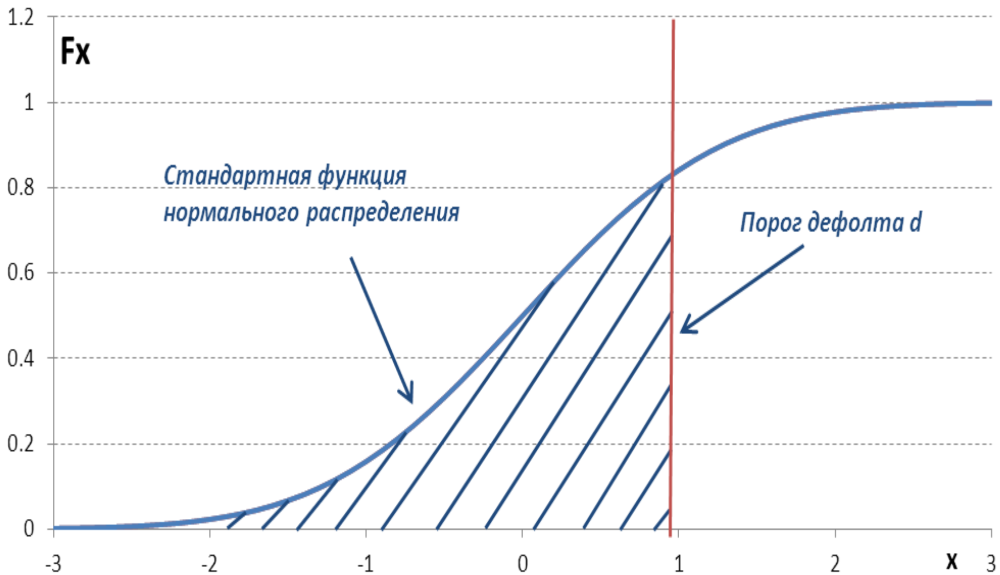
Основываясь на решении уравнения, вероятность того, что активы пересекут обязательства можно представить в следующем виде:

$$P(A_{iT} < D_i) = P[d_i + z_i < 0] = \Phi(z_i \le d_i)$$

Ф – стандартный нормальный закон распределения.

© ВТБ 2013

Модель кредитного портфеля. (Распределение Васичека.)



Пусть ***ψ*** стандартно нормально распределенная величина N(0,1). Вероятность того, что Винеровский процесс пересечет обязательства D равна, вероятности того, что ***ψ≤d***, где ***d*** функция от горизонта моделирования, начальных активов и обязательств.

Представим ***ψ*** в следующем виде ***ψ=(ρ^0.5)*v+((1- ρ)^0.5)*w***, где ***v,w*** – стандартно нормально распределенные величины.

Пусть ***w*** отвечает за риск конкретного заемщика, а ***v*** за глобальный риск, одинаковый для всех заемщиков. Для каждого момента времени мы получаем тот же Винеровский процесс для конкретного контрагента, но связанные процессы для нескольких контрагентов.

$$P[\sqrt{\rho}\cdot f+\sqrt{1-\rho}\cdot\varepsilon_i<d_i]=P\left(\varepsilon_i\le\frac{d_i-\sqrt{\rho}\cdot f}{\sqrt{1-\rho}}\right)=\hat{O}\left(\frac{d_i-\sqrt{\rho}\cdot f}{\sqrt{1-\rho}}\right)$$

Оценка параметра d может быть получена для каждой однородной группы исходя из наблюдаемой частоты дефолтов и биноминального распределения

$$L=\binom{n}{k}p^k(1-p)^{n-k}=\binom{n}{k}P[\psi\le d]^k(1-P[\psi\le d])^{n-k}\Rightarrow\widehat{d}=\Phi^{-1}(\widehat{p})$$

Если в портфеле ***R*** рейтинговых групп, то вероятность наблюдать ***x*** дефолтов может быть выражена следующим образом:

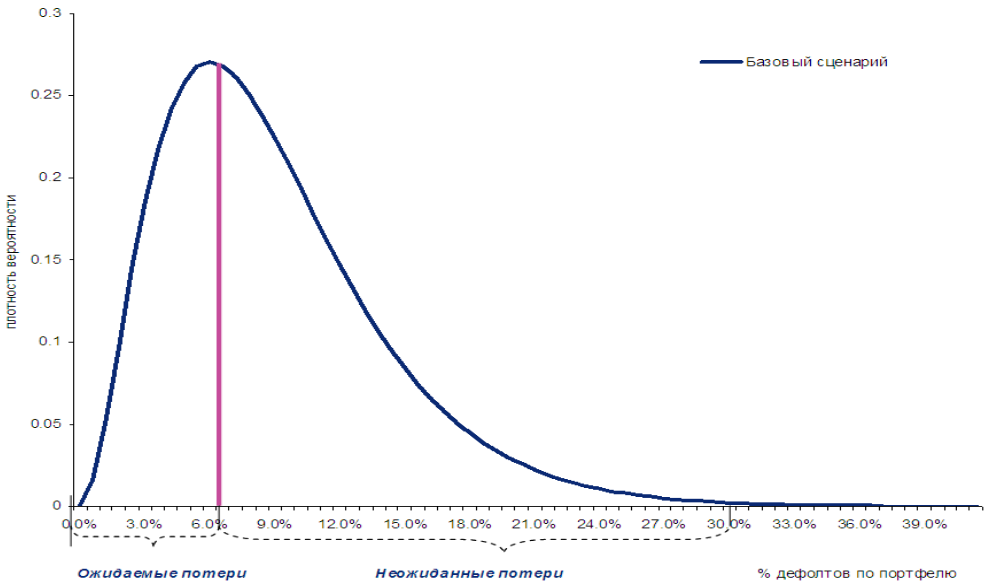
$$P\left[L=\sum_{j=1}^{R}x_j\right]=\int_{-\infty}^{+\infty}\prod_{j=1}^{R}\binom{n_r}{x_r}\cdot\left(\Phi\left(-\frac{\Phi^{-1}(\widehat{p}_r)+\sqrt{\rho_j}\cdot f}{\sqrt{1-\rho_j}}\right)\right)^{x_j}\cdot\left(1-\Phi\left(-\frac{\Phi^{-1}(\widehat{p}_r)+\sqrt{\rho_j}\cdot f}{\sqrt{1-\rho_j}}\right)\right)^{n_j-x_j}\cdot\Phi(f)df$$

© ВТБ 2013

Модель кредитного портфеля. (Распределение Васичека.)

Функция распределения Васичека имеет следующий асимптотический вид :

$$f_{\infty} = \sqrt{\frac{\rho}{1-\rho}} \exp\left(-\frac{1}{2\rho} \left(\sqrt{1-\rho}\, N^{-1}(\theta) - N^{-1}(p) \right)^2 + \frac{1}{2} (N^{-1}(\theta))^2 \right)$$



Для получения оценок распределения Васичека с помощью метода максимального правдоподобия необходимы следующие входные параметры за рассматриваемый период:

- Количество дефолтов по портфелю
- Количество заемщиков в портфеле
- Рейтинговая структура портфеля

Для реализации модели KMV нет необходимости в использовании трудно извлекаемых данных таких как: динамика активов компании, динамика капитализации компании, рыночная стоимость облигаций и т.д. Такой подход важен для России, где рыночные котировки нельзя рассматривать как надежный индикатор оценок для определения корреляции.
Данное обстоятельство является весомым преимуществом данной модели перед другими моделями кредитного риска портфеля.

© ВТБ 2013

План презентации

1. Введение. Принципы моделей кредитного портфеля.
2. Модель кредитного портфеля (распределение Васичека).
3. **Модификация распределения потерь и стресс-тест.**
4. Модификация Многофакторной модели распределения потерь.
5. Сравнение результатов калибровки с IRB формулой Базель-2.
6. Аспекты практической имплементации модели.
7. Заключение.

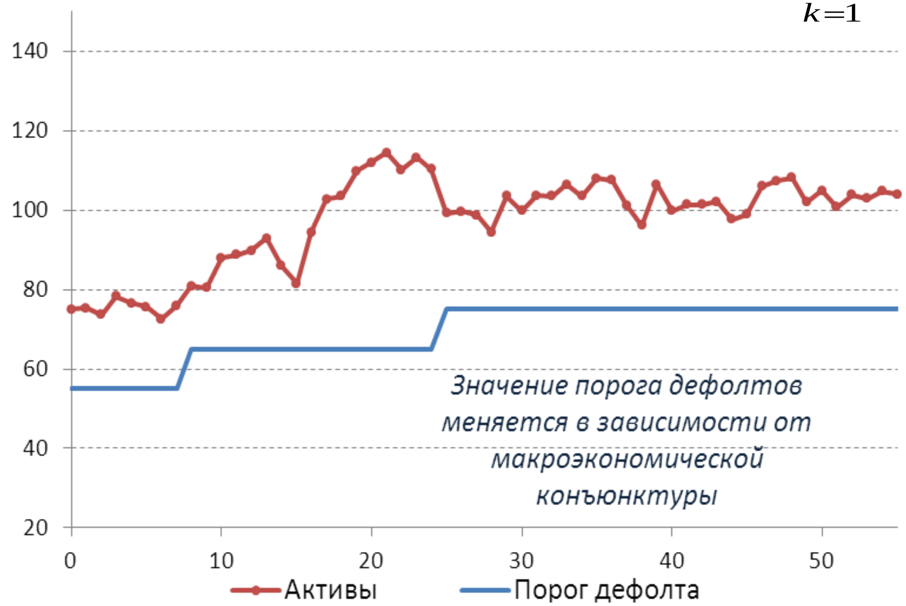
© ВТБ 2013

Модификация распределения потерь по портфелю.

В модель Васичека могут быть включены драйверы макроэкономического риска. Связь между наблюдаемыми риск-факторами и вероятностью дефолта устанавливается через линейную модель порога. Таким образом для порога ***d*** предполагается следующая линейная спецификация:

$$d_i = \alpha_0 + \sum_{k=1}^{n} \alpha_k M_{kt}$$

,где ***M*** представляет макроэкономические факторы, ***α*** **–** постоянные коэффициенты



Вероятность дефолта примет следующий вид:

$$P\left(\varepsilon_i \le \frac{\alpha_0 + \sum_{k=1}^{p} \alpha_k \cdot M_{kt} - \sqrt{\rho} \cdot f}{\sqrt{1-\rho}}\right) = \hat{O}\left(\frac{\alpha_0 + \sum_{k=1}^{p} \alpha_k \cdot M_{kt} - \sqrt{\rho} \cdot f}{\sqrt{1-\rho}}\right)$$

Функция правдоподобия для ***T*** периодов и ***R*** рейтинговых классов будет иметь следующий вид:

$$L(k) = \prod_{t=1}^{T} \binom{n}{x} \cdot \int_0^1 \prod_{j=1}^{r} \left(N\left(-\frac{\alpha_{0j} + \sum_{i=1}^{p} \alpha_{ij} M_i + \sqrt{\rho_j} \cdot N^{-1}(f)}{\sqrt{1-\rho_j}} \right) \right)^{x_r} \cdot \left(1 - N\left(-\frac{\alpha_{0j} + \sum_{i=1}^{p} \alpha_{ij} M_i + \sqrt{\rho_j} \cdot N^{-1}(f)}{\sqrt{1-\rho_j}} \right) \right)^{n_r - x_r} dN(f)$$

© ВТБ 2013

Модель кредитного портфеля.

Учитывая влияние макроэкономических показателей на распределение дефолтов, приходим к следующему алгоритму модели кредитного портфеля:

I. ***Определение набора влияющих макроэкономических показателей*** → *Выбираем макроэкономические показатели, имеющие наибольшую статистическую связь с частотой дефолтов по портфелю.*

II. ***Построение предиктивных моделей для каждого показателя, вошедшего в модель*** → *Для каждого макроэкономические показателя, вошедшего в модель, строим статистическую модель, позволяющую делать прогноз на период моделирования.*

III. ***Генерация сценариев для каждого показателя*** → *Для каждого сценария метода Монте-Карло прогнозируем значения макроэкономического показателя, на основе статистической модели.*

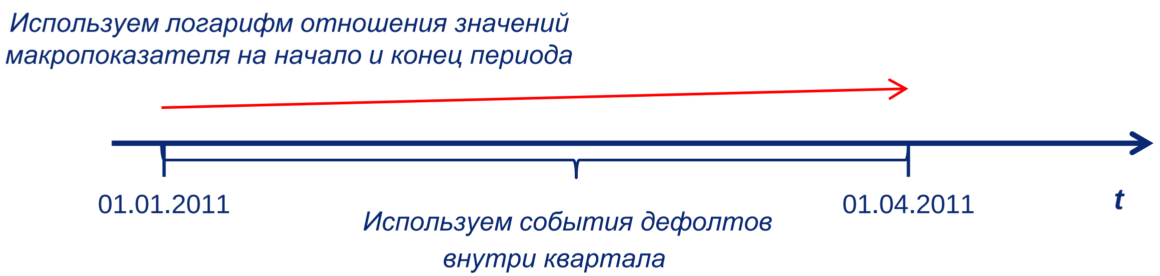
IV. ***Построение функции плотности распределения потерь*** → *С помощью метода Монте-Карло строим распределение потерь по портфелю.*

© ВТБ 2013

Модель кредитного портфеля. Выбор макроэкономических показателей.

Макроэкономические показатели входящие в модель выбираются исходя из следующих параметров:

1. Временного лага, относительно частоты дефолтов.



2. Корреляции между показателем и частотой дефолтов.

ln(x(t)/x(t-1))	*t*	*t-1*	*t-2*	*t-3*
Реальный ВВП	-0.79	-0.85	-061	-053
Доходы населения	-0.69	-0.73	-0.53	-0.15
Нефть	-0.81	-0.79	-075	-0.30
РТС	-0.75	-0.63	-0.5	-0.49

Решение о включении показателя в модель принимается на основе теста отношения правдоподобия (LR-test).

3. Экономического смысла влияния показателя на вероятность дефолта.

Знак значения корреляции и постоянного коэффициента в модели должны иметь экономический смысл.

© ВТБ 2013

Модель кредитного портфеля. Матрицы миграций

Существует два подхода к оценке матриц миграций:

1. Оценка матриц миграций методом когорт (cohort method):

Вероятность перехода из рейтинга *i* в рейтинг *j* определяется соотношением:

$$P_{ij} = \frac{N_{ij}(t)}{N_i(t)}$$

N_{ij} - Количество миграций из рейтинга *i* в рейтинг *j* за период t.

N_i - Количество контрагентов с рейтингом *i* на начало периода.

Можно считать матрицы миграций отдельно для периодов роста экономики и спада.

2. Описание миграций рейтингов Марковским процессом:

Миграция рейтингов представляется в виде неоднородного Марковского процесса с конечномерным пространством возможных состояний. Матрица миграций определяется через интенсивности миграций рейтингов:

$$P(t) = \exp(\Lambda t) \qquad \Lambda = \{\lambda_{ij}\} \qquad i, j = 1,.., R$$

Оценка интенсивности может быть произведена с помощью регрессии Кокса (Cox regression):

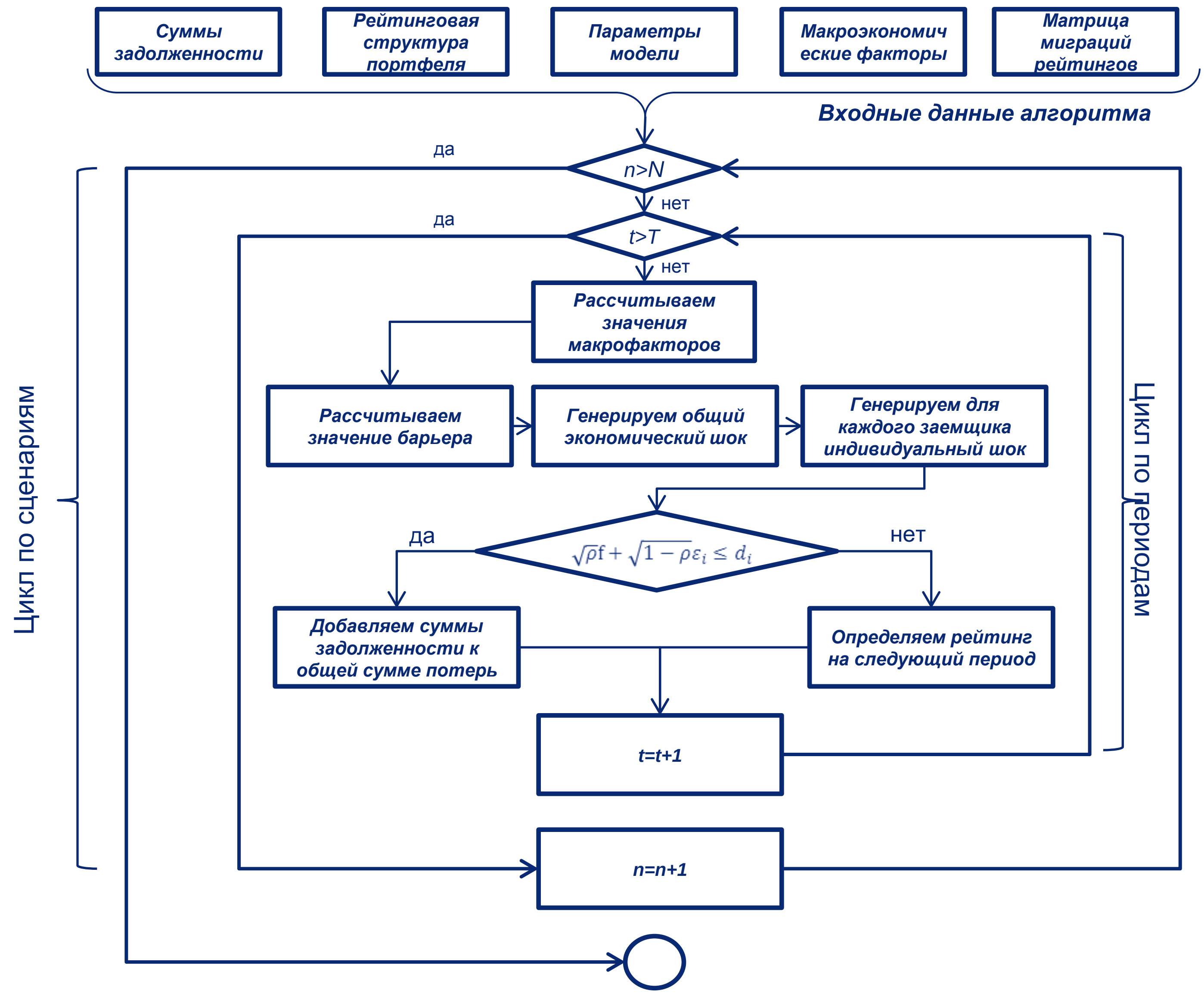
$$\lambda_{hji} = Y_{hi}(t)\alpha_{hji}(t, Z_i(t)) \qquad \alpha_{hji}(t, Z_i(t)) = \alpha_{hjo}(t)\exp(\beta' Z_{hji}(t))$$

Y_{hi} – индикаторная переменная равная 1, если заемщик i находился в состоянии h в момент времени t

Более подробно: Lando, D. and T. Skodeberg, 2002, “Analyzing Ratings Transitions and Rating Drift with Continuous Observations”

© ВТБ 2013

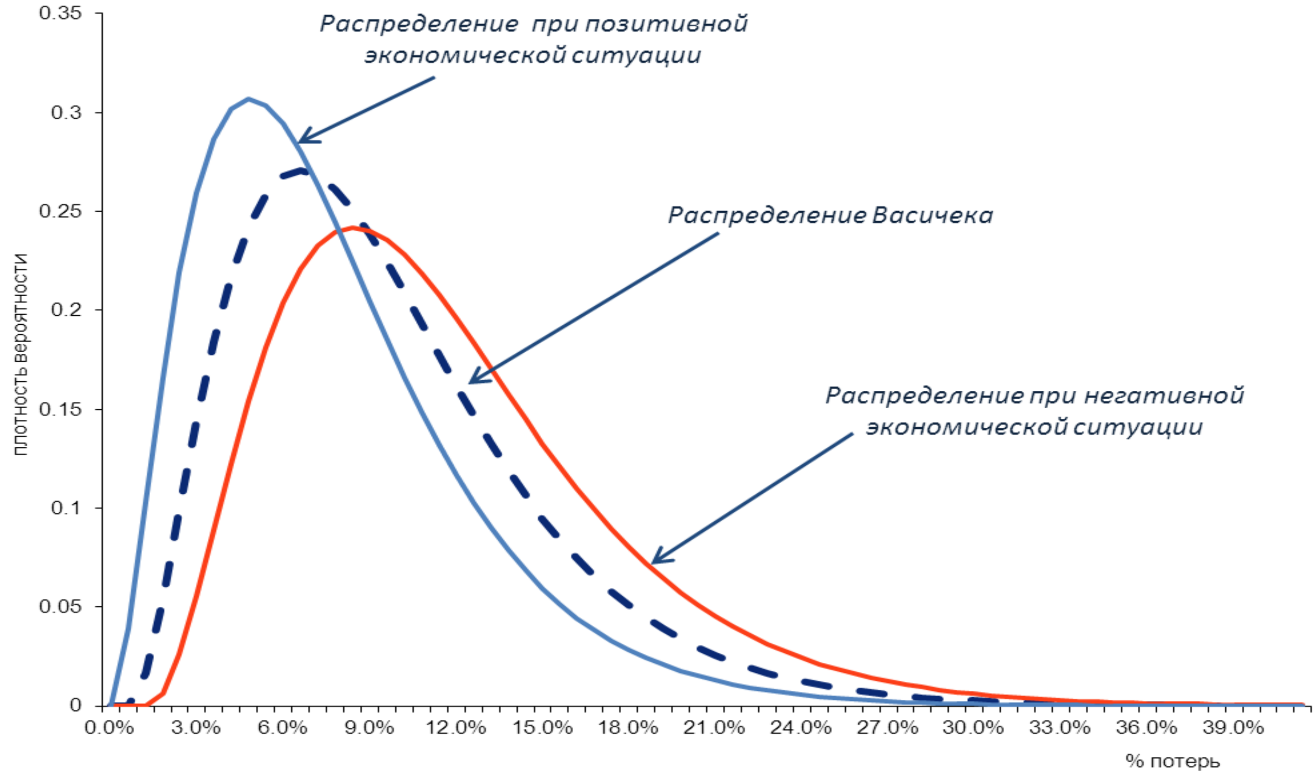
Модель кредитного портфеля. Алгоритм расчета функции распределения потерь.



© ВТБ 2013

Модель кредитного портфеля. Результат работы модели.

В результате работы алгоритма получаем функцию распределения потерь. Получаемая функция чувствительная к текущей экономической ситуации. Таким образом размер Экономического капитала также меняется в зависимости от текущей экономической ситуации.



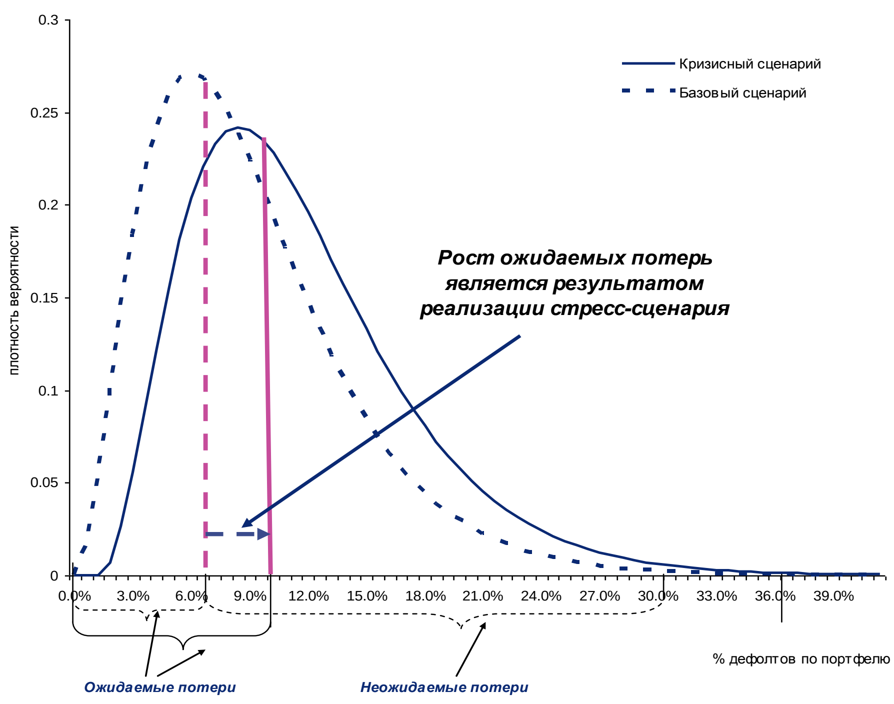
Средним вариантом является ситуация, когда модель кредитного портфеля строится на основе данных о дефолтах и корреляции между ними. Эту модель можно назвать нейтральной моделью (нейтральной к макроэкономическим изменениям в ходе кредитного цикла). Введение в модель макроэкономических параметров позволяет оценить оптимистическое и пессимистическое распределение потерь в следующем году на основе различных макроэкономических сценариев. Такие кривые потерь будут соответственно находиться левее и правее нейтральной модели.

© ВТБ 2013

Стресс-тест (модификация модели KMV)

С точки зрения регулятора для поддержания стабильности финансовой системы модели кредитного риска должны быть дополнены концепцией стресс-тестирования

“ 726. Банки должны показать, что намеченные ими внутренние ориентиры по капиталу достаточно обоснованы и отвечают общему характеру риска и операционной среды. При оценке достаточности капитала руководству банка необходимо учитывать конкретную стадию цикла деловой активности. Необходимо проводить тщательное, ориентированное на будущее стресс-тестирование для выявления возможных событий или изменений в рыночной конъюнктуре, способных неблагоприятно повлиять на банк. Несомненно, руководство банка несет главную ответственность за наличие у банка соразмерного его рискам капитала […]”



Включая в модель стрессовые значения макроэкономических параметров можно получить оценку возможных потерь по кредитному портфелю в случае реализации неблагоприятных макроэкономических прогнозов.

Построение кривой распределения потерь в ситуации стресс- сценария позволит определить две новые характеристики кредитного портфеля после «шока»:

1.Новая (стрессовая) величина вероятных потерь, что позволит оценить рост резервов Банка;

2. Новую величину экономического капитала, который необходим банку, чтобы покрыть увеличившуюся величину неожидаемых потерь.

© ВТБ 2013

Пояснения и выводы

1. Если мы моделируем потери по портфелю на один базовый период (квартал или год), то макроэкономический показатель моделируется, как случайная величина, параметры которой зависят от текущих значений показателя. Таким образом распределение потерь зависит от текущий экономической ситуации. Обобщение на большие количество периодов происходит таким же образом, значение показателя будет зависеть от предыдущих значений.

2. Включение в модель данных о поведении макроэкономических факторов позволяет расширить подход к модели кредитного портфеля, сделав его более прикладным, и отвечать на вопрос – каковы будут потери при определенных макроэкономических параметрах. Эти параметры, могут носить шоковый характер, а могут быть обычными в рамках макроэкономичекой тенденции.

3. На наш взгляд для целей построения системы планирования и распределения капитала между подразделениями банка при планировании на 1 год более корректно использовать модель портфеля, включающую в себя макроэкономические показатели. При разработки стратегического плана кредитной организации – от 3 до 5 лет, более корректно использовать модель кредитного портфеля, основанной только на среднециклических данных о дефолтах, без включения макроэкономических параметров.

© ВТБ 2013

План презентации

1. Введение. Принципы моделей кредитного портфеля.

2. Модель кредитного портфеля (распределение Васичека).

3. Модификация распределения потерь и стресс-тест.

4. **Модификация Многофакторной модели распределения потерь.**

5. Сравнение результатов калибровки с IRB формулой Базель-2.

6. Аспекты практической имплементации модели.

7. Заключение.

© ВТБ 2013

Многофакторная модель.

Многофакторная модель является обобщением однофакторной модели. При спецификации многофакторной модели предполагается наличие ***K*** факторов представляющие например отрасли или регионы. При этом предполагается наличие корреляции между этими факторами.

$$R_{it}^1 = \sqrt{\rho_1}\, f_t^1 + \sqrt{1-\rho_1}\,\varepsilon_{it}^1$$

$$\dots$$

$$R_{it}^K = \sqrt{\rho_K}\, f_t^K + \sqrt{1-\rho_K}\,\varepsilon_{it}^K$$

$$r_{ij} = corr(f^i, f^j)$$

$$p_i^1(f_t^1) = \Phi\left(\frac{T_1 - \sqrt{\rho_1}\, f_t^1}{\sqrt{1-\rho_1}}\right)$$

$$\dots$$

$$p_i^1(f_t^1) = \Phi\left(\frac{T_K - \sqrt{\rho_K}\, f_t^K}{\sqrt{1-\rho_K}}\right)$$

,где $T_1 \dots T_K$ представляют барьеры и также могут быть представлены как линейные функции от макроэкономических показателей.

Вследствие независимости событий дефолта при реализации фактора ***f*** **,** можно применить «закон больших чисел» и предположить что доля дефолтов по категории пропорциональна вероятности дефолта:

$$P[X = n_1 p_1(y_1) + \dots + n_K p_K(y_K) \mid f_1 = y_1, \dots, f_2 = y_2] = 1$$

$$P[X \le x] = E[P[X \le x \mid f_1 = y_1, \dots, f_2 = y_2]] = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} P[X \le x \mid f_1 = y_1, \dots, f_2 = y_2]\,\varphi(f_t^1, \dots, f_t^K)\, df_t^1 \dots df_t^K$$

Логарифмическая функция правдоподобия :

$$l = \sum_{t=1}^{T} \ln\Big(\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \binom{n_t^1}{s_1} p^1(f_t)^{s_1}(1-p^1(f_t))^{n_t^1 - s_1} \cdot \binom{n_t^2}{s_2} p^2(f_t)^{s_2}(1-p^2(f_t))^{n_t^2 - s_2} \cdot \dots \cdot \binom{n_t^2}{s_2} p^K(f_t)^{s_K}(1-p^K(f_t))^{n_t^K - s_K} \boxed{\varphi(f_t^1, \dots, f_t^K)}\, df_t^1 \dots df_t^K\Big)$$

Многомерная функция нормального распределения с ковариационной матрицей Σ.

Требуется оценка дополнительно ***0,5*K(K-1)*** *параметров.*

© ВТБ 2013

Модификация Многофакторной модели распределения потерь

Модель представляет модификацию модели Васичека, и позволяет учитывать корреляцию как внутри отраслей (между рейтинговыми классами) так и между отраслями.

Модификация модели Васичека заключается в том, что вводится корреляция между системными факторами каждой отрасли. Таким образом корреляция между дефолтами различается не только по группам рейтингов, но и по отраслям. Описанный эффект достигается, тем что Винеровский процесс определяющий изменение стоимости активов во времени, можно представить следующим образом:

$$z_{i,k} = \varepsilon_i \cdot \sqrt{1-\rho} + y_k \cdot \sqrt{\rho \cdot (1-\rho_{\text{int}})} + z \cdot \sqrt{\rho \cdot \rho_{\text{int}}}$$

$$P(x = 1 | y_k = \tau, z = \upsilon) = N\left(\frac{N^{-1}(PD) - \sqrt{\rho \cdot (1-\rho_{\text{int}})}\,\tau - \sqrt{\rho \cdot \rho_{\text{int}}} \cdot \upsilon}{\sqrt{1-\rho}} \right)$$

$\varepsilon_i, y_k, z -$ стандартные нормальные распределения; $i, k -$ номер рейтинга и отрасли соответственно;

$\rho, \rho_{\text{int}} -$ Корреляция между изменением стоимости активов между рейтингами и отраслями соответственно.

Корреляция между отраслевыми шоками передается через общий фактор z, характеризует общее состояние экономики.

Логарифмическая функция правдоподобия:

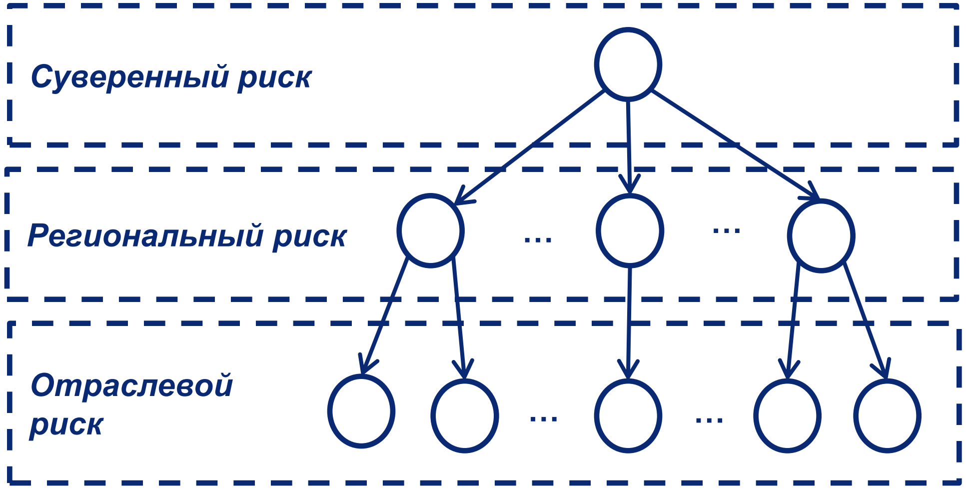
$$l = \sum_{t=1}^{T} \ln\Big(\int_{-\infty}^{+\infty} \Big[\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \binom{n_t^2}{s_2} p^2(f_t)^{s_2} (1-p^2(f_t))^{n_t^2 - s_2} \cdot \ldots \cdot \binom{n_t^2}{s_2} p^K(f_t)^{s_K} (1-p^K(f_t))^{n_t^K - s_K} \varphi(f_t^1, \ldots, f_t^K) \cdot \varphi(y_1) \cdot \ldots \cdot \varphi(y_k)\, dy_1 \ldots dy_k \Big] \varphi(z)\, dz$$

Переменная Z определяет состояние экономики, а переменная $\rho \cdot \rho_{\text{int}}$ определяет степень влияния изменений в экономике на отдельную отрасль

© ВТБ 2013

Модификация Многофакторной модели распределения потерь

Многофакторная модель предложенного вида может быть обобщена для возможности включения в модель большего количества групп риска. Представим стандартную случайную величину в следующем виде: ***z=a1*f1+a2*f2+…+an*fn+b*eps,*** где ***f1,..,fn*** – системные факторы разного уровня, ***eps***-индивидуальный риск. ***f1, .. ,fn, eps*** – имеют стандартное нормальное распределение.



Коэффициенты ***a1, … , an, b*** выбираются таким образом, чтобы с их сумма была равна 1.

Вероятность дефолта по рейтинговой группе ***I*** будет иметь следующий вид:

$$p_i^1(f_1^T,\dots,f_n^T)=\Phi\left(\frac{T_i-\sum_{i=1}^{n}a_i\cdot f_i^T}{b}\right)$$

Таким образом в модель может быть включена корреляция по странам, регионам, отраслям и так далее. При этом корреляция внутри каждой категории риска будет различна.

© ВТБ 2013

План презентации

1. *Введение. Принципы моделей кредитного портфеля.*
2. *Модель кредитного портфеля (распределение Васичека).*
3. *Модификация распределения потерь и стресс-тест.*
4. *Модификация Многофакторной модели распределения потерь.*
5. ***Сравнение результатов калибровки с IRB формулой Базель-2.***
6. *Аспекты практической имплементации модели.*
7. *Заключение.*

© ВТБ 2013

Однофакторная модель.

Частным случаем многофакторной модели является однофакторная модель, асимптотическая версия которой используется в рамках IRB подходов соглашения Basel II. Таким образом, сделав ряд упрощений в разработанных процедурах калибровки, мы можем проверить сопоставимость корреляций реально наблюдаемых в портфеле с корреляциями IRB Basel II:

$$R_{it}^{1} = \sqrt{\rho_1}\, f_t^{1} + \sqrt{1-\rho_1}\,\varepsilon_{it}^{1}$$

$$\ldots$$

$$R_{it}^{K} = \sqrt{\rho_K}\, f_t^{K} + \sqrt{1-\rho_K}\,\varepsilon_{it}^{K}$$

$$r_{ij} = corr(f^{i}, f^{j}) == 1$$

$$\Rightarrow$$

$$p_i^1(f_t) = \Phi\left(\frac{T_1 - \sqrt{\rho_1}\, f_t}{\sqrt{1-\rho_1}}\right)$$

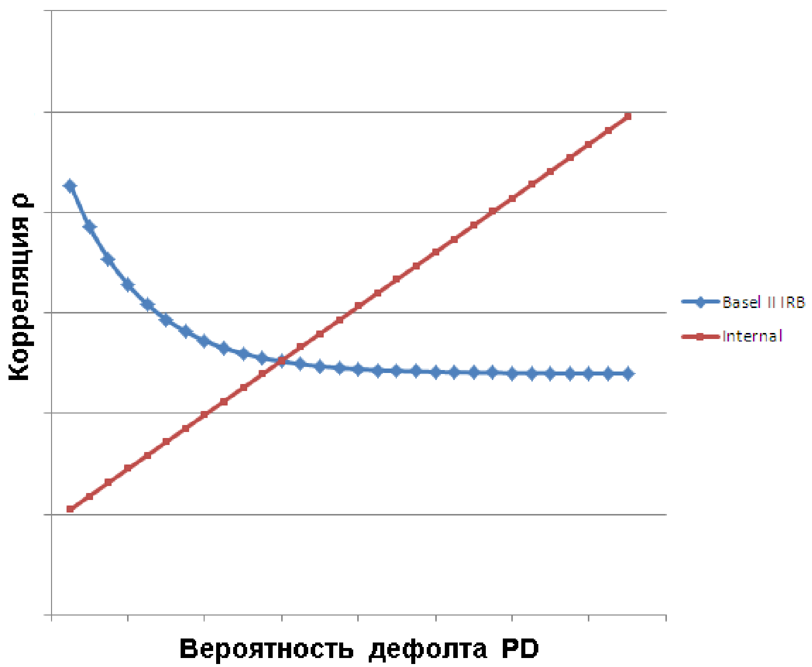
$$\ldots$$

$$p_i^1(f_t) = \Phi\left(\frac{T_K - \sqrt{\rho_K}\, f_t}{\sqrt{1-\rho_K}}\right)$$

Полученные значения корреляции в зависимости от PD достаточно точно аппроксимируются линейной функцией.

Зависимость корреляции от вероятности дефолта получилась обратной относительно калибровки IRB Basel II – заемщики с высокими PD характеризуются более высокими показателя корреляции. Одновременно, для низкодефолтных рейтинговых классов калибровка IRB Basel II переоценивает значения корреляций.

Потенциально, выявленная закономерность может быть объяснена гипотезой о высоком уровне государственной поддержки крупных компаний в период кризиса (по сравнению с ЕС), которая «сглаживает» влияние системного фактора на данные компании (т.е. снижает корреляцию).



План презентации

1. Введение. Принципы моделей кредитного портфеля.
2. Модель кредитного портфеля (распределение Васичека).
3. Модификация распределения потерь и стресс-тест.
4. Модификация Многофакторной модели распределения потерь.
5. Сравнение результатов калибровки с IRB формулой Базель-2.
6. **Аспекты практической имплементации модели.**
7. Заключение.

© ВТБ 2013

Калибровка логарифмической функции правдоподобия

$$l=\sum_{t=1}^{T}\ln\Big(\int_{-\infty}^{+\infty}\Big[\int_{-\infty}^{+\infty}\dots\int_{-\infty}^{+\infty}\binom{n_t^2}{s_2}p^2(f_t)^{s_2}(1-p^2(f_t))^{n_t^2-s_2}\cdot\ldots\cdot\binom{n_t^2}{s_2}p^K(f_t)^{s_K}(1-p^K(f_t))^{n_t^K-s_K}\varphi(f_t^1,\ldots,f_t^K)\cdot\varphi(y_1)\cdot\ldots\cdot\varphi(y_k)dy_1\ldots dy_k\Big]\varphi(z)dz$$

Согласно предпосылкам, заложенным в модель, количество дефолтов подчиняется биномиальному закону распределения. В результате, попытка численно оценить значение логарифмической функции правдоподобия в заданной точке, даже для относительно небольших портфелей с числом заёмщиков ~1000, ведет к выходу численных значений за границы стандартных для ПО чисел двойной точности (double-precision). Возможны два пути решения данной проблемы:

1) Применение асимптотической аппроксимации биномиального распределения (например, пуассоновским распределением). Данный подход существенно снижает вычислительную нагрузку. В связи с чем, это единственно возможное решение при необходимости множественной калибровки моделей, например, в случае применения подхода кросс-валидации.
2) Применение специализированных пакетов с неограниченной точностью вычислений (ограничение накладывается исключительно размером ОЗУ сервера).

В связи с тем, что портфельные модели и модели стресс-тестирования становятся частью регулятивного процесса, корректность асимптотической калибровки, на наш взгляд, обязательно должна быть проверена «точным» решением с использованием специализированного ПО.

Возможным решением является использование свободно распространяемого ПО «R» – интерпретируемого высокоуровневого языка программирования, позволяющего одновременно использовать концепции объектно-ориентированного программирования и матричную арифметику (как в Matlab), благодаря чему реализованный «точный» алгоритм калибровки:

- Поддерживает неограниченную точность вычислений (использовался пакет «Rmpfr»)
- Обеспечивает высокую скорость расчетов путем:
  1) использования исключительно векторной арифметики без применения циклов;
  2) распараллеливания наиболее ресурсоемких вычислений (численное интегрирование) (использовался пакет «parallel»)

Высокий уровень корреляций между макропоказателями

Негативным свойством макропоказателей РФ является их высокая взаимная корреляция. На этапе отбора макропоказателей при калибровке модели стресс-тестирования для стабилизации модели следует отсечь показатели с высокой взаимной корреляцией (как правило, порог составляет 75%). Однако, данный подход не универсален:

1) В соответствии с требованиями (рекомендациями) регулятивных органов, сценарий стресс-тестирования может включать информацию о изменении высоко коррелированных показателей.

2) Макро-аналитики способны прогнозировать ограниченный перечень показателей.

3) Корреляция показателей может со временем изменяется (см. пример изменения корреляции цены нефти Urals и цены на кв. м. в г. Москва). Соответственно, при отсутствии механизмов «подавления» корреляции, может потребоваться регулярный пересмотр факторов в модели, что нарушит консистентность стресс-тестирования во времени и существенно увеличит трудозатраты.

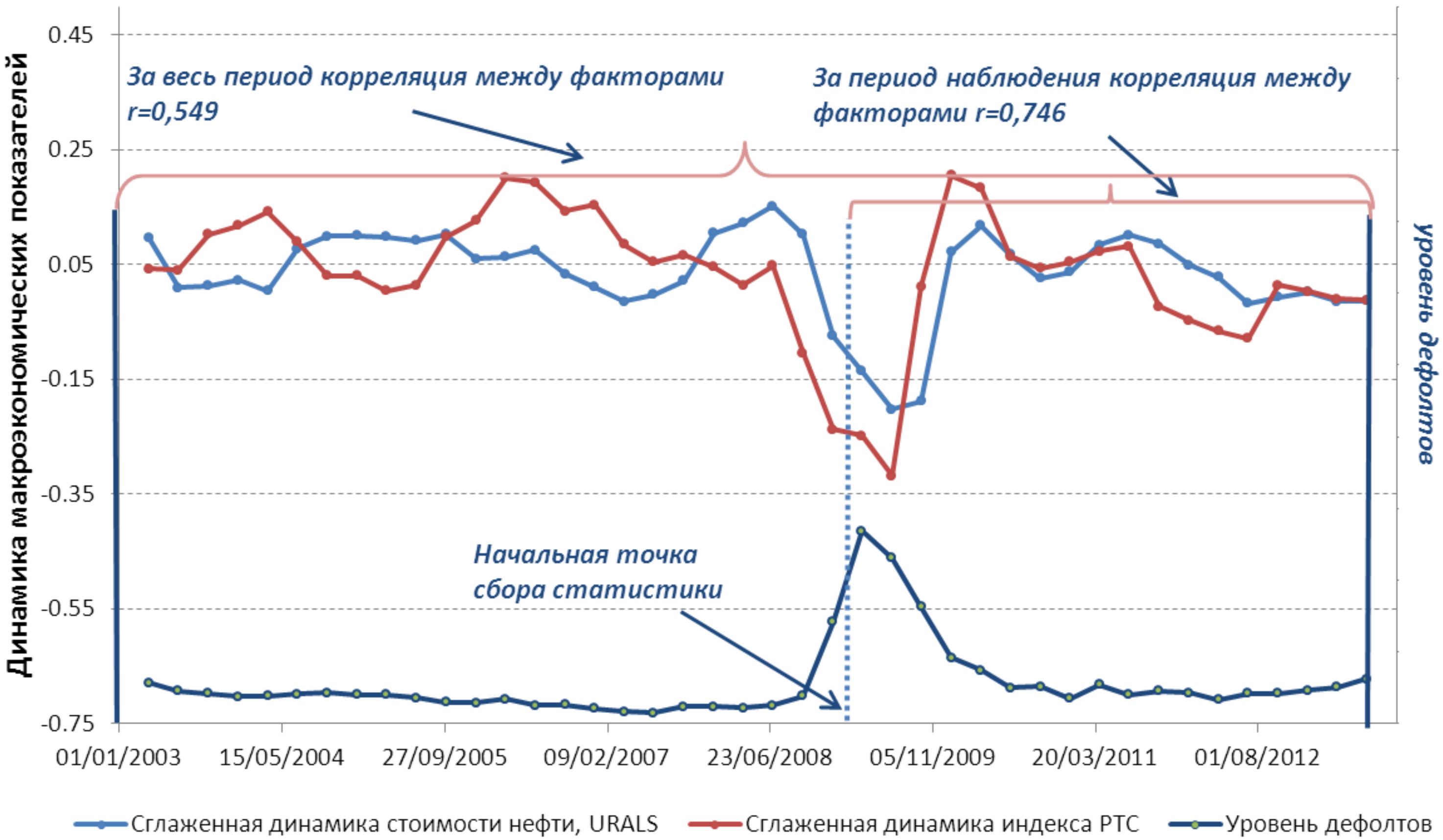
Существует два основных способа «подавления» корреляции:

PCA регрессия – модель калибруется на «главных компонентах», которые являются линейными комбинациями исходных переменных, при этом, линейная комбинация строится таким образом, чтобы «главные компоненты» были независимы (имели единичную матрицу корреляций).

Регуляризация – метод, изменяющий процесс калибровки модели путем добавления специализированной «штрафной» компоненты к функции максимального правдоподобия.

Мультиколлинеарность. Причины.

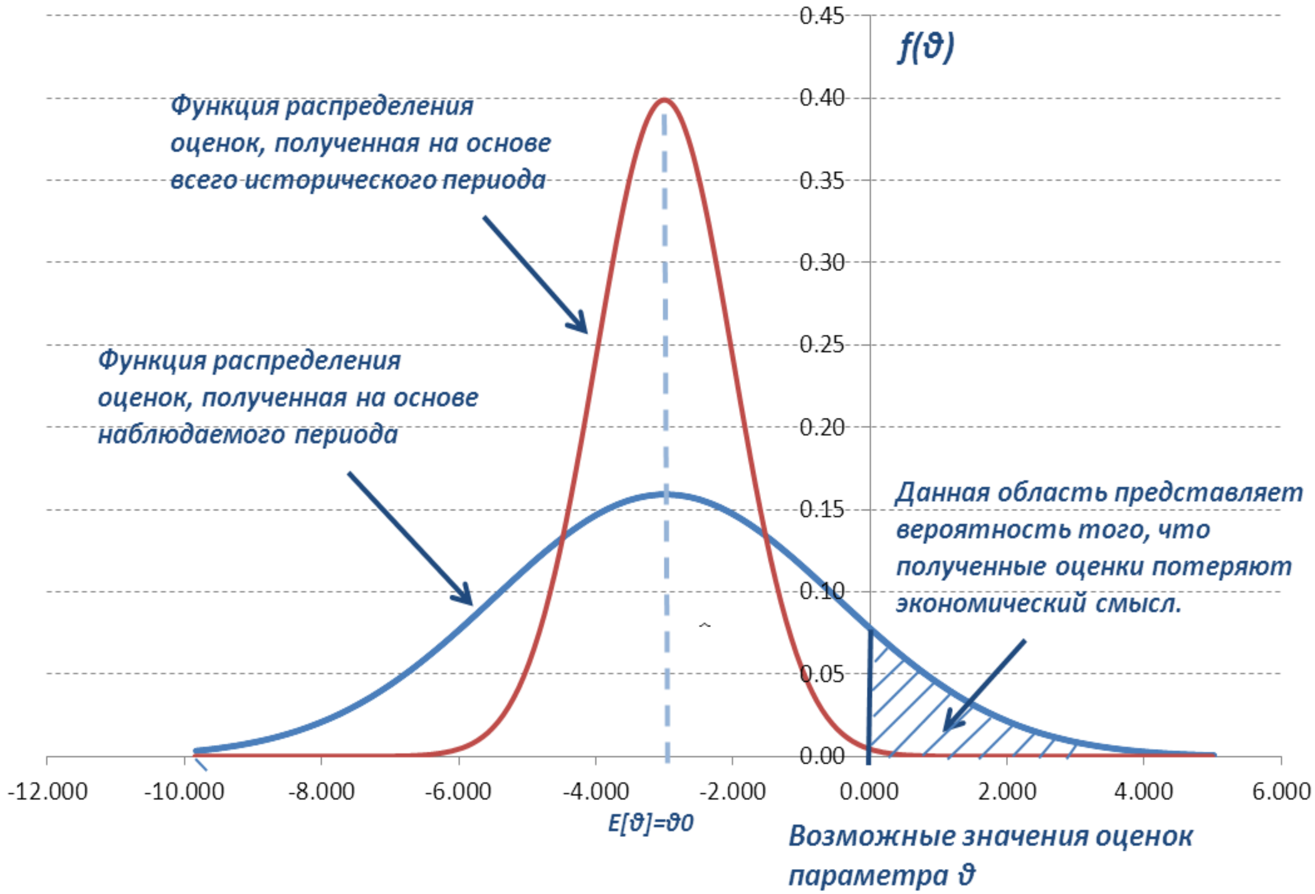
Мультиколлинеарность возникает в случаях существования достаточно тесных линейных статистических связей между объясняющими переменными.



Таким образом мультиколлинеарность может возникнуть в результате недостаточного объема статистики.

© ВТБ 2013

Мультиколлинеарность. Причины.

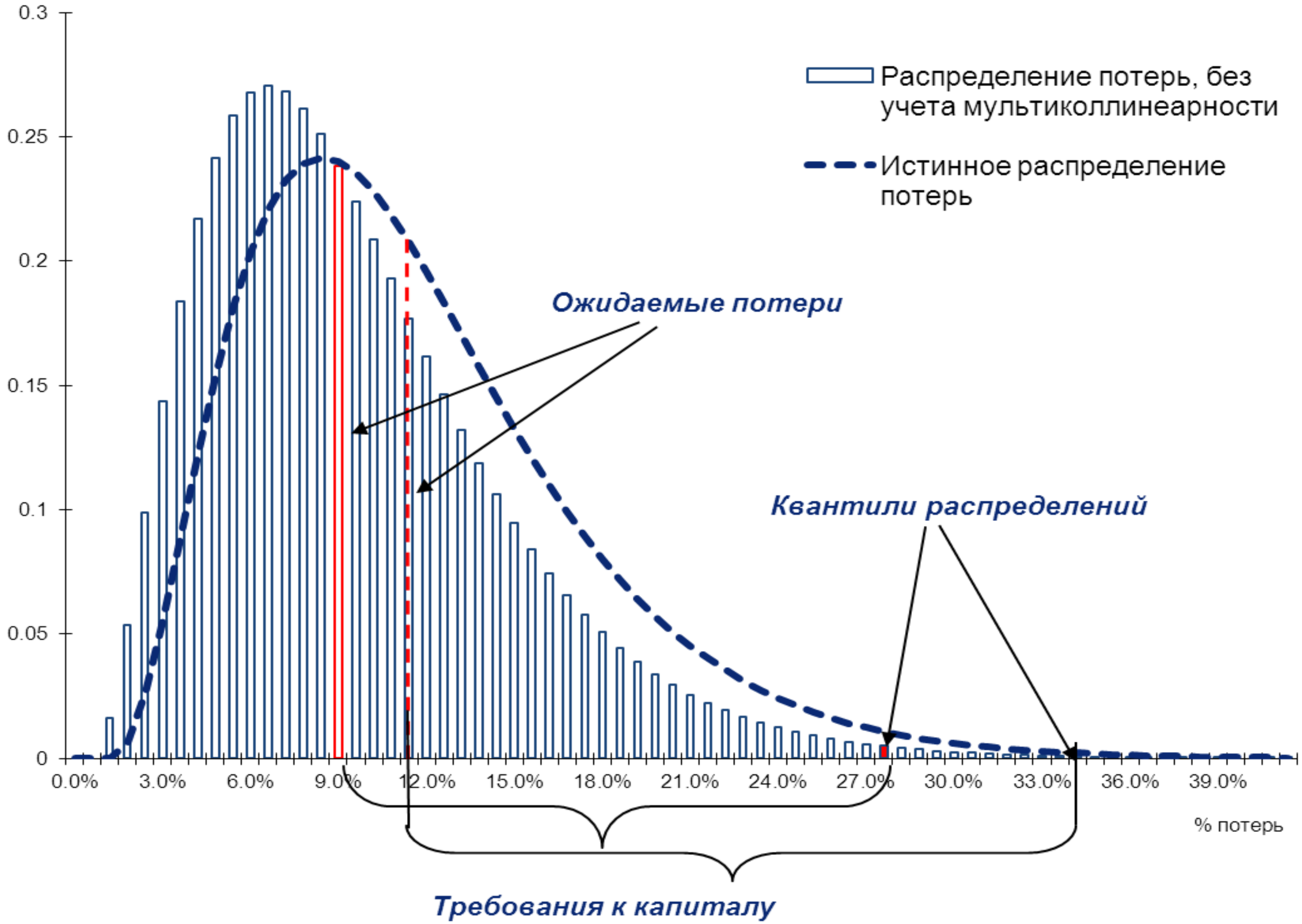


Наличие мультиколлинеарности приводит:

1. К увеличению ошибок получаемых оценок
2. Возникновению вероятности того, что получаемые оценки потеряют экономический смысл.

© ВТБ 2013

Мультиколлинеарность. Влияние на результат стресс-теста.



Использование неточных оценок может приводить к значительным искажениям результата стресс-теста. При этом результат может быть как и завышенным, так и заниженным.

© ВТБ 2013

Мультиколлинеарность. Признаки.

$$L(\alpha)=\prod_{t=1}^{T}\binom{n}{x}\cdot\int_{0}^{1}\prod_{j=1}^{r}\left(N\left(-\frac{\alpha_{0j}+\sum_{i=1}^{p}\alpha_{ij}M_i+\sqrt{\rho_j}\cdot N^{-1}(f)}{\sqrt{1-\rho_j}}\right)\right)^{x_r}\cdot\left(1-N\left(-\frac{\alpha_{0j}+\sum_{i=1}^{p}\alpha_{ij}M_i+\sqrt{\rho_j}\cdot N^{-1}(f)}{\sqrt{1-\rho_j}}\right)\right)^{n_r-x_r}dN(f)$$

При получении оценок параметров функции распределения можно выделить следующие признаки, свидетельствующие о влиянии мультиколлинеарности на оценки параметров:

1. Отношение дисперсии полученных оценок $\sigma^2(\alpha_i)$ к полученной оценке α_i больше 0,7. Высокая ошибка свидетельствует о нестабильности полученных оценок параметров распределения .

2. Число обусловленности Гессиана функции правдоподобия значительно больше 1, т.е.

$$cond(H(\alpha))=\|H(\alpha)\|\cdot\|H^{-1}(\alpha)\|=\frac{\max|\lambda_i|}{\min|\lambda_i|}>>1,$$ λ_i – собственные числа Гессиана.

Учитывая, что $\sigma^2=H(\theta)^{-1}\cdot I(\theta)\cdot H(\theta)^{-1}$, то чем хуже будет обусловлена матрица ***H***, тем больше будет дисперсия получаемых оценок. Кроме того плохая обусловленность Гессиана может свидетельствовать о том, что минимизирующая последовательность не сходится.

3. Знак перед полученной оценкой не имеет экономического смысла.

© ВТБ 2013

Мультиколлинеарность. Методы устранения.

1.Метод главных компонент (PCA).

Согласно этому методу начальные ***p*** макрофакторов трансформируются в ***k*** новых переменных, между которыми отсутствует корреляция и называемых главными компонентами. Фактически производится замена переменных и согласно определенному правилу отбрасываются менее значимые переменные. Коэффициенты перед новыми переменными получаются с помощью метода максимального правдоподобия. Полученные оценки не будут подвержены влиянию эффекта мультиколлинеарности.

2. Метод регуляризации Тихонова.

Данный метод предложен советским математиком А.Н.Тихоновым для решения некорректных задач. Вместо максимизации исходной функции правдоподобия, будем искать минимум следующего функционала:

$$l^{\lambda}(\alpha,\rho)=-\ln(L(\alpha_0,\alpha,\rho))+\lambda\cdot\Omega(\alpha)$$, где

$\Omega(\alpha)-$ неотрицательная, возрастающая функция.

$\lambda-$ неотрицательный параметр регуляризации

Слагаемое $\lambda\cdot\Omega(\alpha)$ штрафует за слишком большие значения ***α***, и таким образом контролирует дисперсию оценок.

© ВТБ 2013

Метод главных компонент (1).

В модели связь между наблюдаемыми риск-факторами и вероятностью дефолта устанавливается через линейную модель порога. Таким образом для порога ***d*** предполагается следующая линейная спецификация:

$$d_i = \alpha_0 + \sum_{k=1}^{p} \alpha_k m_{kt} = \alpha_0 + M \cdot \alpha$$

,где ***M*** представляет макроэкономические факторы, ***α*** – постоянные коэффициенты

I. ***Нормируем макроэкономические переменные***

Преобразуем макроэкономические переменные следующим образом: $m_i^0 = (m_i - E[m_i]) \cdot D[m_i]^{-0.5}, i = 1, \ldots, p$

II. ***Рассчитываем V - матрицу собственных векторов матрицы C.***

,где ***C*** *корреляционная матрица нормированных макроэкономических переменных.* $V \cdot V' = I$

III. ***Производим замену переменных и выбираем главные компоненты.***

$$d_i = \alpha_0 + M^0 \cdot \alpha = \alpha_0 + M \cdot V \cdot V' \cdot \alpha = \alpha_0 + Z \cdot \beta,$$

$$Z = M \cdot V, \qquad \beta = V' \cdot \alpha$$

IV. ***Получаем оценки ММП***

$$\hat{\beta} = \arg\max(L(\alpha_0, \beta, \rho))$$

© ВТБ 2013

Метод главных компонент (2).

1. Выбор значимых компонент.

Одной из возможных стратегий выбора главных компонент может быть следующая: основываясь на том, что чем меньше собственное значение соответствующее компоненте, тем менее она информативна, таким образом исключаем компоненты до тех пор, пока оставшиеся компоненты описывают не менее 85% исходной дисперсии. Таким образом выбираем ***r*** наиболее значимых компонент, для которых выполняется следующее неравенство:

$$\frac{\sum_{j=1}^{r} \lambda_j}{p} > 0.85$$

После исключения неинформативных компонент, уравнение будет иметь следующий вид:

$d_i = \alpha_0 + Z_r \cdot \beta_r$, $Z_r = M \cdot V_r$, $V_r -$ матрица , полученная из матрицы собственных векторов путем удаления столбцов, соответствующих исключенным компонентам.

2. Переход к коэффициентам перед макроэкономическими показателями.

$\alpha^c = V_r \cdot \beta_r -$ в результате обратного преобразования получаем вектор коэффициентов перед центрированными макрофакторами.

Для перехода к коэффициентам перед ненормированными переменными следует провести следующие преобразования:

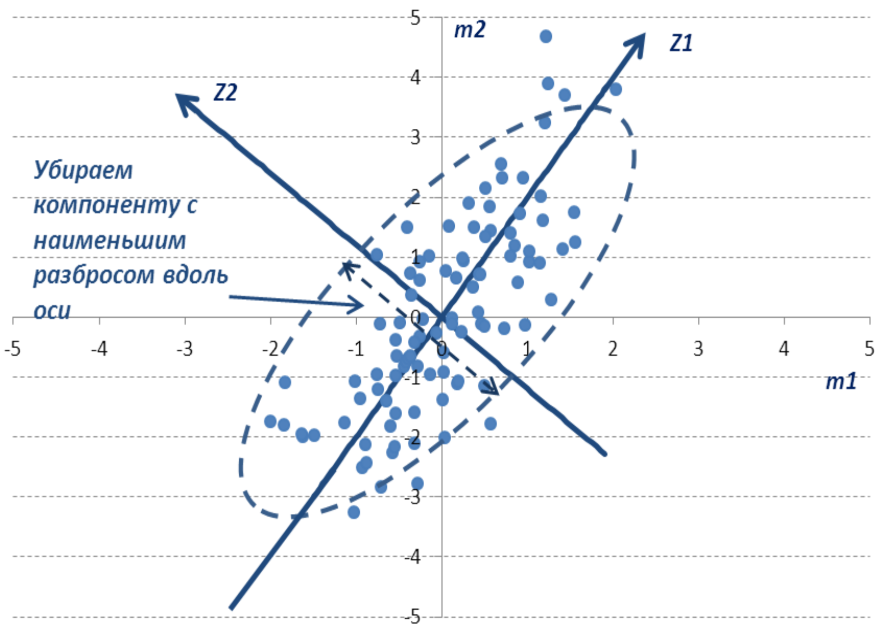
$$\alpha_i = \frac{\alpha_i^c}{s_i}, i = 1,\dots, p \qquad \alpha_o = \alpha_o^c - \frac{\alpha_1^c \cdot \overline{m}_1}{s_1} - \frac{\alpha_2^c \cdot \overline{m}_2}{s_2} - \dots - \frac{\alpha_p^c \cdot \overline{m}_p}{s_p}$$

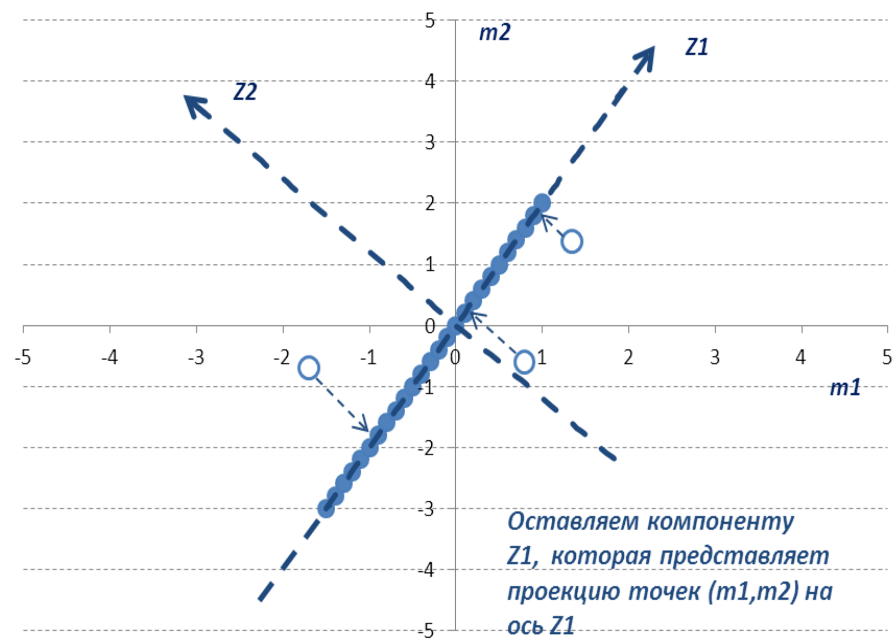
© ВТБ 2013

Метод главных компонент (3). Геометрическая интерпретация.

Рассмотрим следующую задачу поиска двух оптимальных параметров ***a1** и **a2.*** Тогда порог дефолтов будет определяться следующим уравнением: $d_i = a1 \cdot m1 + a2 \cdot m2$

Перейдем к новой системе координатам (Z1,Z2), в которой точки становятся независимы.





Чтобы уменьшить дисперсию модели, убираем компоненту, которая обладает меньшей дисперсией.

$d = \alpha_0 + Z_2 \cdot \beta_2, \quad Z_2 = (m1, m2) \cdot V_r$, V_r – матрица , полученная из матрицы собственных векторов путем удаления столбца, соответствующего первой компоненте.

Регуляризация.

$$L(x) = \prod_{t=1}^{T} \binom{n}{x} \cdot \int_0^1 \prod_{j=1}^{r} \left(N\left(-\frac{\alpha^*{}_{0j} + \sum_{i=1}^{p} \alpha^*{}_{ij} M^0{}_i + \sqrt{\rho_j} \cdot N^{-1}(f)}{\sqrt{1-\rho_j}} \right) \right)^{x_r} \cdot \left(1 - N\left(-\frac{\alpha^*{}_{0j} + \sum_{i=1}^{p} \alpha^*{}_{ij} M^0{}_i + \sqrt{\rho_j} \cdot N^{-1}(f)}{\sqrt{1-\rho_j}} \right) \right)^{n_r - x_r} dN(f)$$

Пусть $\widetilde{z} = (\widetilde{\alpha}^*{}_0, \widetilde{\alpha}^*, \widetilde{\rho})$ – набор истинных параметров, при которых ***L(x)*** достигает максимума.

Рассмотрим максимизацию ***L(x)*** как некорректно поставленную задачу, т.е. незначительные изменениях в обучающей выборке, приводят к значительным изменениям получаемых оценок (эффект мультиколинеарности). Тогда оценки, полученные при минимизации следующего функционала

$$z = (\alpha_0^*, \alpha^*, \rho) = agr\min(-\ln(L(\alpha_0^*, \alpha^*, \rho)) + \lambda \cdot \sum_{i=1}^{p} \alpha_i^{*2}) = agr\min(\Theta)$$

будут обладать следующим свойством:

Для любого ***ε>0*** существует такое ***λ***, что для набора параметров ***z*** минимизирующего ***Θ*** выполняется неравенство*: $\rho(z, \widetilde{z}) \le \varepsilon$

$\rho(z, \widetilde{z})$ – Евклидово расстояние между наборами параметров.

Таким образом, полученные регуляризованные оценки могут быть как угодно близки к истинным оценкам, и не будут зависеть от малых изменений обучающей выборки.

*** *Доказательство данного неравенства см. Тихонов А.Н., Арсенин В.Я. «Методы решения некорректных задач».***

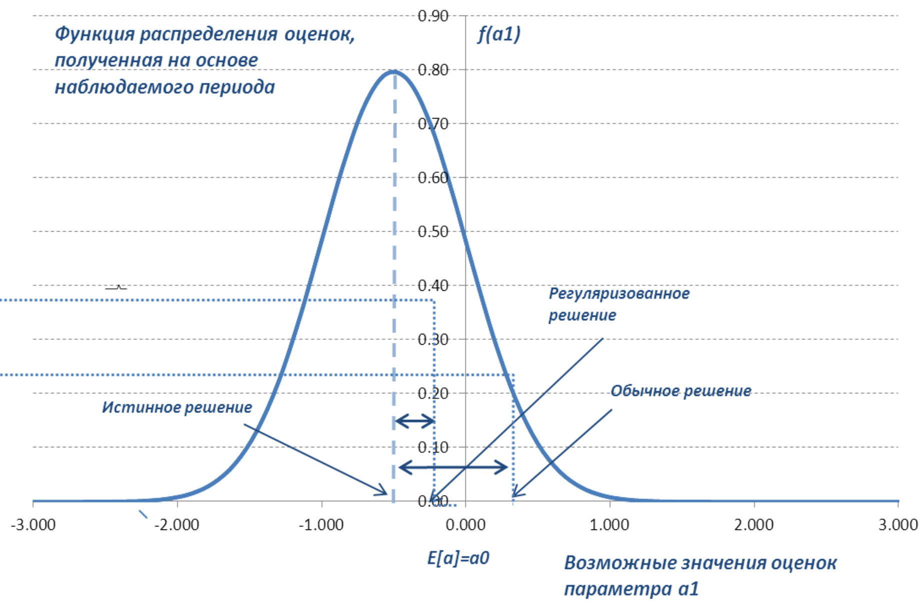
© ВТБ 2013

Регуляризация. Геометрическая интерпретация.

Рассмотрим следующую задачу поиска двух оптимальных параметров ***a1** и **a2.*** Тогда порог дефолтов будет определяться следующим уравнением: $d_i = a1 \cdot m1 + a2 \cdot m2$

Оптимальные параметры функции распределения ***L(x)*** , будем искать минимизируя следующий функционал:

$$l^{\lambda}(\alpha, \rho) = -\ln(L(d(a1, a2), \rho)) + \lambda \cdot (a1^2 + a2^2)$$

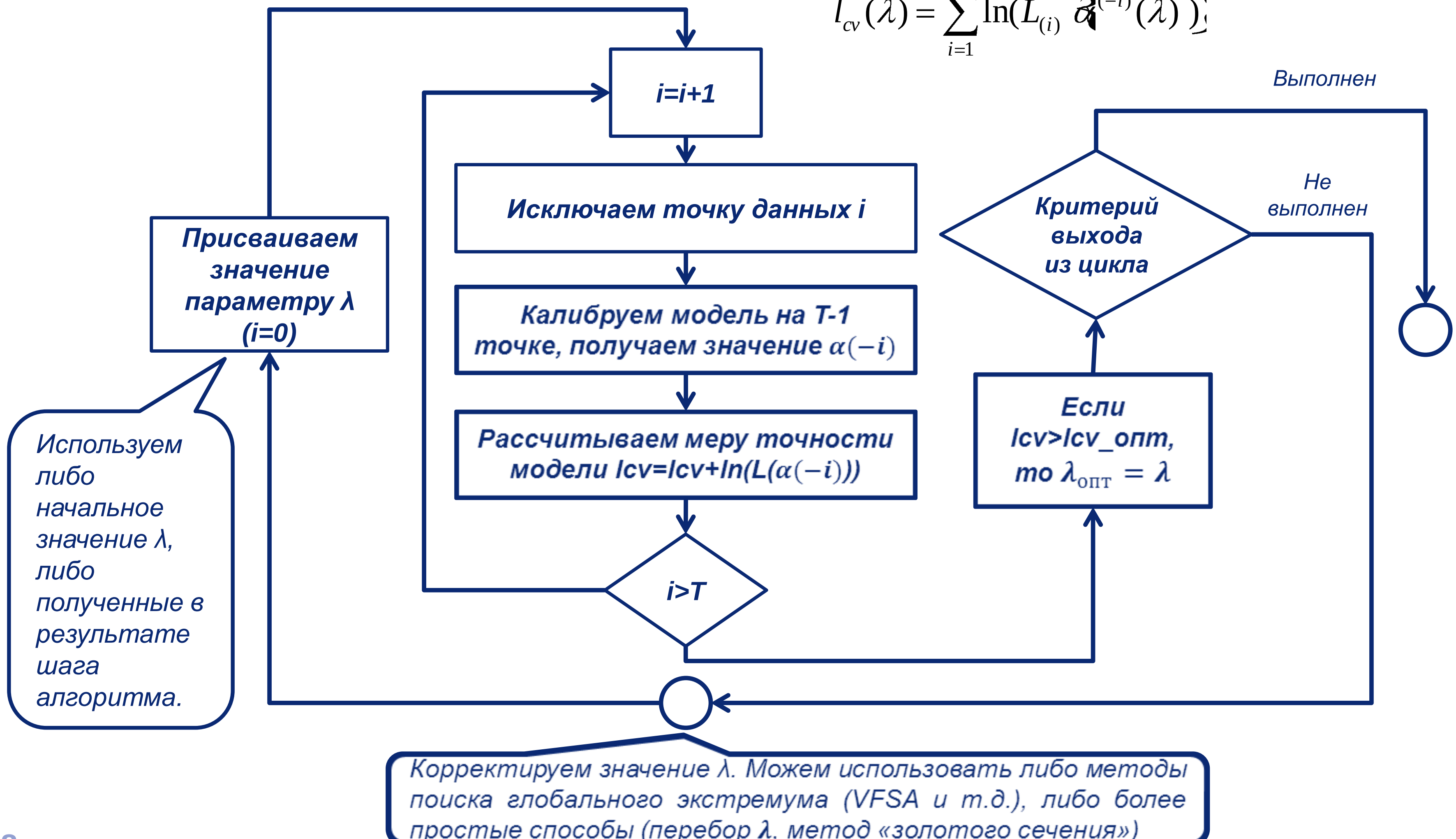


© ВТБ 2013

Регуляризация. Выбор оптимального регуляризирующего параметра.

Для сравнения различных параметров λ в качестве меры будем использовать значение функции правдоподобия на тестовой выборке. Тестовая выборка формируется на основе one-leave-out кросс-валидации. Выбор оптимального значения λ производится согласно следующему алгоритму:

$$l_{cv}(\lambda)=\sum_{i=1}^{T}\ln\left(L_{(i)}\left(\alpha^{(-i)}(\lambda)\right)\right)$$



© ВТБ 2013

Сравнение методов.

	Метод главных компонент	*Регуляризация*
Преимущества	1. Интуитивно понятный алгоритм. 2. Сравнительно небольшая скорость обучения алгоритма.	1. Гарантирует сходимость к оптимальному решению. 2. Оценки устойчивы к малым изменениям в выборке. 3. Помогает избежать переобучение модели.
Недостатки	1. При исключении одной из компонент теряется часть информации.	1. Полученные оценки являются смещенными. 2. Слишком большое время обучения.

© ВТБ 2013

План презентации

1. Введение. Принципы моделей кредитного портфеля.

2. Модель кредитного портфеля (распределение Васичека).

3. Модификация распределения потерь и стресс-тест.

4. Модификация Многофакторной модели распределения потерь.

5. Сравнение результатов калибровки с IRB формулой Базель-2.

6. Аспекты практической имплементации модели.

7. Заключение.

© ВТБ 2013

Заключение

В ходе презентации представлена модель расчета кредитного риска по портфелю активов. Данная модель основана на распределение дефолтов с учетом корреляции (распределение Васичека) и обладает рядом преимуществ:

1.Учитывает корреляцию между дефолтами контрагентов.

2.Для оценки параметров не требует знания ненаблюдаемых переменных.

3.Учитывает влияние экономической ситуации на распределение дефолтов.

4.Достаточно просто поддается модификации.

5.Интуитивно понятна.

Кроме того данная модель может использоваться для проведения стресс-тестирования кредитного портфеля.

Однако у предложенной модели есть следующие недостатки:

1. Является моделью построенной по принципу «сверху-вниз», поэтому возникают проблемы с учетом специфического риска заемщика. Соответственно больше применима для однородных портфелей, таких как : портфель розничных кредитов или портфелей кредитов малому и среднему бизнесу.
2. Требует классификации заемщиков в однородные группы (согласно вероятности дефолта), что также приводит к усреднению результатов.

© ВТБ 2013

Спасибо за внимание!

ВТБ