



AI 周观察

数据专题研究
证券研究报告

国金证券研究所

分析师：刘道明（执业 S1130520020004）
liudaoming@gjzq.com.cn分析师：黄晓军（执业 S1130525080005）
huangxiaojun@gjzq.com.cn

Claude Fable5 发布后被禁，Oracle 正面临毛利率下滑

摘要

- 本周国内外大模型动态频发。海外市场中，ChatGPT、Claude 和 Gemini 活跃度小幅上升。模型层面，Anthropic 虽面向公众推出具备分级过滤机制的 Claude Fable5 模型，但随后因美国政府发布涉及国家安全的出口管制指令而被迫对所有用户全面停服召回；谷歌则发布了实验性开源语言模型 DiffusionGemma，首次将扩散机制引入文本生成，凭借非线性“全块感知”能力在代码填空等结构化任务中展现出独特优势。国内市场，多数聊天类应用网页端活跃度保持稳定，小米技术团队则正式开源了终端 AI 编程助手 MiMo Code V0.1.0，主打“模型 Agent 协同优化”与持久记忆系统，在多项权威测试集上表现优异，并支持全语音操控。
- Oracle 2026 年第一季度（FY2026Q4）收入达到 192 亿美元，同比增长 21%，创单季历史新高，主要由 OCI 云基础设施和云应用业务共同驱动。其中云基础设施收入同比增长 93%，反映 AI 训练与推理工作负载和数据库服务需求强劲。云应用收入 41 亿美元，同比增长 10%，SaaS 递延收入增长 16%，显示传统企业应用仍保持稳健增长。全年收入首次突破 670 亿美元，非 GAAP 营业利润达到 290 亿美元，同比增长 16%，非 GAAP EPS 为 7.63 美元。全年毛利率下降约 5 个百分点，2027 财年毛利率预计继续下降。

风险提示

芯片制程发展与良率不及预期
中美科技领域政策恶化
智能手机销量不及预期



内容目录

海外市场行情回顾.....	3
Claude Fable5 发布后被禁，Google 发布 DiffusionGemma	4
Oracle：毛利率正在降低，仍需关注数据中心建设情况.....	5
风险提示.....	7



海外市场行情回顾

图表1: 截至6月12日海外AI相关个股行情

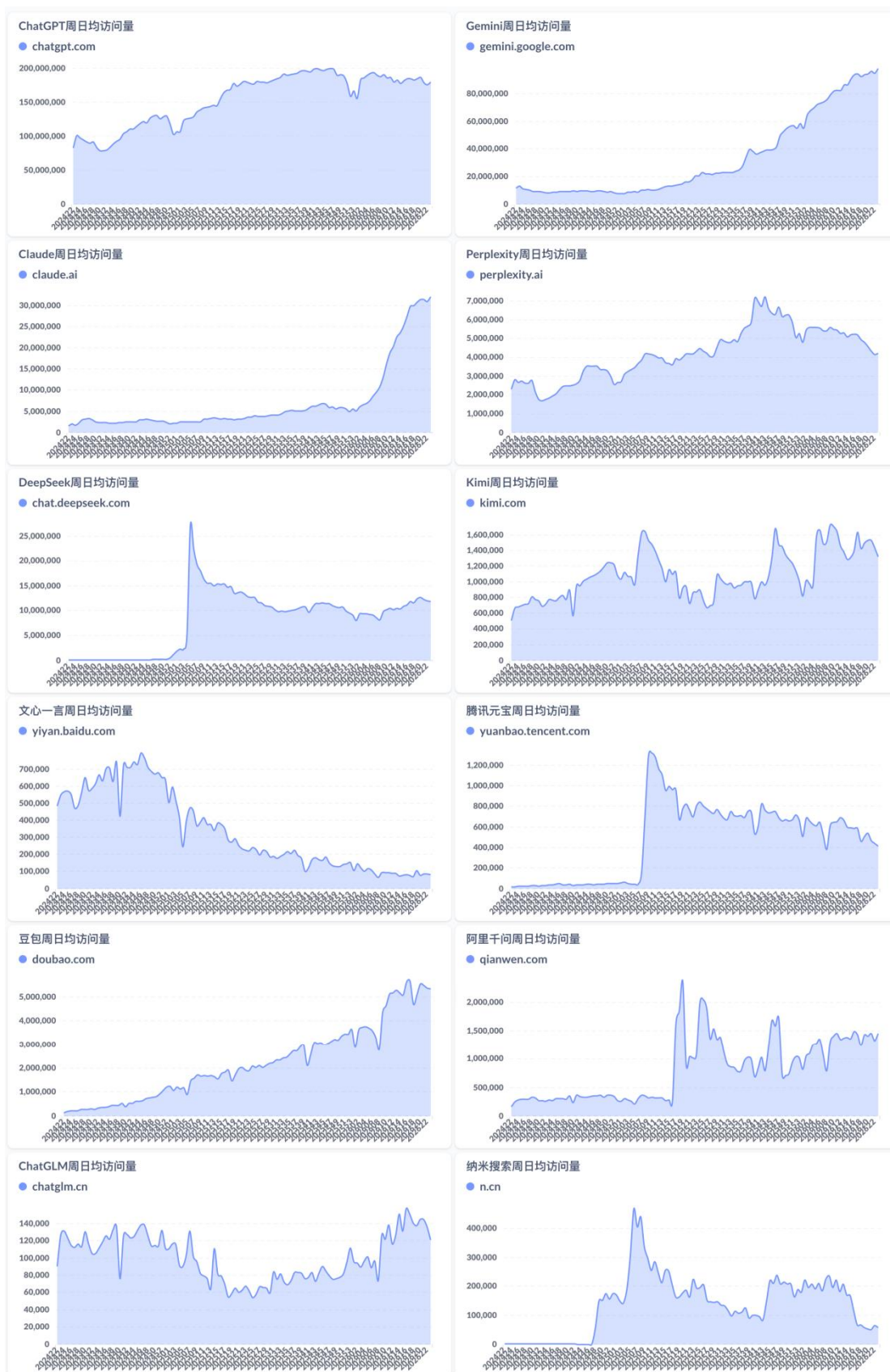
英特尔	124.57	25.61%	芯片设计与制造
超威半导体	511.57	9.69%	芯片设计
美满电子	279.7	6.16%	芯片设计与制造、交换机
C3.AI	10.9	4.51%	云服务
特斯拉	406.43	3.95%	整车、大模型
Palo Alto Networks	279.62	2.78%	网络安全
CrowdStrike	682.8	1.76%	网络安全
Skyworks	73.97	0.54%	芯片设计与制造
Qorvo	98.59	0.32%	芯片设计与制造
戴尔	395.57	0.3%	消费电子产品、服务器、交换机、网络安全
英伟达	205.19	0.04%	芯片设计、交换机
博通	382.07	-0.95%	芯片设计与制造、交换机
Zscaler	129.52	-0.96%	网络安全
Innodata	100.9	-1.05%	云服务
HP	25.24	-1.33%	消费电子产品
DataDog	229.9	-1.8%	芯片设计
高通	211.72	-1.95%	芯片设计
Mongodb	342.8	-2.26%	云服务
Snowflake	232.78	-2.3%	云服务、软件
谷歌	359.68	-2.4%	云服务、大模型、广告
亚马逊	238.55	-3.04%	IaaS、电商平台、自营硬件、物流配送、广告变现
Meta	566.98	-4.39%	消费电子产品、大模型、广告
苹果	291.13	-5.27%	消费电子产品
Palantir	127.99	-5.56%	云服务
微软	390.74	-6.22%	消费电子产品、云服务、软件
Zeta Global	20.21	-8.22%	软件
Cloudflare	228.48	-8.65%	网络安全
Salesforce	165.89	-10.65%	云服务
Gitlab	27.79	-10.7%	云服务

来源: Reuters、国金证券研究所



Claude Fable5 发布后被禁，Google 发布 DiffusionGemma

图表2：聊天助手类 AI 应用活跃度



来源：Similarweb、国金证券研究所

本周，海外聊天助手类应用活跃度中，ChatGPT、Claude 和 Gemini 小幅上升。国内多数聊



天类应用网页端活跃度较为稳定。

国外模型方面, Anthropic 于周二面向公众推出了新一代 Claude Mythos 系列中的 Fable5 模型。该模型此前因网络安全隐患被限制测试, 本次发布采取了分级过滤机制, 涉及敏感领域的查询会被自动重定向至低阶模型。伴随大众版发布, Anthropic 向 Project Glasswing 计划中的签约企业与网络安全伙伴提供了无限制的高阶版本, 目前该合作网络已覆盖 15 个以上国家的约 200 家组织。鉴于其极强的关键基础设施漏洞识别能力, 白宫已建立常态化机制在发布前对其进行安全审查。由于 Fable5 收费标准比以往翻番, 且每月需向 xAI 支付巨额数据中心租赁费用, Anthropic 面临巨大资金压力, 然而在 6 月 11 日下午 5 点 21 分, 美国政府引用国家安全权限发布出口管制指令, 要求暂停所有外国人对 Fable5 和 Mythos5 的访问, 导致 Anthropic 在当晚对所有用户采取全面停服的召回措施。受此影响, 新建会话将使用默认模型或 Opus4.8 运行, 现有 Fable5 会话将报错中止。此外, 谷歌于 2026 年 6 月 10 日正式发布了实验性开源语言模型 DiffusionGemma, 首次将图像 AI 中的扩散机制引入文本生成领域, 打破了传统自回归范式, 可实现一次性并行输出 256 个标记的词块。该模型拥有 260 亿参数, 依托 MoE 架构, 单步激活参数为 38 亿。在英伟达的深度优化下, 其在 H100 显卡上输出速度可达每秒 1000 个 token, 在 RTX5090 等高端消费级显卡上也能突破每秒 700 个 token。虽然其文本生成质量与准确率略逊于传统的 Gemma4 系列, 但其独特的“全块感知”能力使其在文本后补、代码填空、数独求解以及氨基酸序列等非线性、结构化数据处理任务中展现出显著优势, 目前已基于 Apache2.0 协议在 Hugging Face 开源并全面兼容 vLLM、MLX 等框架。

国内模型方面, 小米技术团队正式发布并开源了一款运行在终端的探索性 AI 编程助手 MiMo Code V0.1.0。该项目基于开源项目 OpenCode 二次开发, 采用宽松的 MIT 协议开源。作为一款主打“模型 Agent 协同优化”的自进化工具, MiMo Code 独创了由独立 Subagent 负责的项目记忆、会话检查点和任务进度三重持久记忆系统, 并内置/dream 命令实现记忆沉淀; 同时打造了专属 Harness 系统与 Compose 模式, 支持全流程自动化编码与验证。在相同 MiMo 模型前提下, 单凭 Agent 系统协同优势, MiMo Code 在 SWE-Bench Pro 和 Terminal Bench2 上的得分分别达到 62% 和 73%, 均领先于 Claude Code。该工具内置限时免费的多模态模型 MiMo-V2.5, 性能比肩 Claude Sonnet4.6, 无需注册即可开箱即用, 同时兼容 DeepSeek、Kimi、GLM 等主流 API, 且全面汉化。此外, 该工具还依托 MiMo-V2.5-ASR 语音识别技术, 支持全语音操控的语音交互。

Oracle: 毛利率正在降低, 仍需关注数据中心建设情况

Oracle 2026 年第一季度 (FY2026Q4) 收入达到 192 亿美元, 同比增长 21%, 创单季历史新高, 主要由 OCI 云基础设施和云应用业务共同驱动。其中云基础设施收入同比增长 93%, 反映 AI 训练与推理工作负载和数据库服务需求强劲。云应用收入 41 亿美元, 同比增长 10%, SaaS 递延收入增长 16%, 显示传统企业应用仍保持稳健增长。全年收入首次突破 670 亿美元, 非 GAAP 营业利润达到 290 亿美元, 同比增长 16%, 非 GAAP EPS 为 7.63 美元, 剔除一次性投资收益后为 6.83 美元。经营现金流达到 320 亿美元, 同比增长 54%, 但由于 AI 数据中心建设加速, 全年资本支出净现金流出达到 480 亿美元。

Oracle 剩余履约义务达到 6380 亿美元, 同比增长 363%, 其中 12% 预计未来 12 个月确认收入, 34% 将在 13 至 36 个月确认, 显示未来收入可见度显著提升。公司本季度签署 670 亿美元 AI 基础设施合同, 其中大部分为客户自带硬件或预付款模式, 使相关合同总额达到 750 亿美元, 且这类合同利润率不低于其他合同。OCI 全年向客户交付超过 1.2GW 容量, 2027 财年第一季度交付量接近 1GW, 几乎相当于过去四个季度总和, 说明 Oracle 的 AI 数据中心建设计划正在逐渐实现。

Oracle 披露 Abilene 园区已交付总容量的 42%, 未来 90 天还将交付 35%。Shackelford、Dona Ana County、Saline 和 Port Washington 等大型站点将在 2027 年陆续开始客户交付。其中 Dona Ana County 采用吉瓦级 Bloom 燃料电池供电设计。Oracle 还表示全球 GPU 利用率达到 97.5%, 说明现有 AI 算力资源需求极强, 即便部分 GPU 续约未完成, 也会迅速转售给其他客户。

云数据库收入同比增长 29%, 其中多云收入同比增长 404%, 订单增长 325%, 说明 Oracle Database 在 OCI、Azure 及其他云环境中的部署正在快速扩张。Oracle 表示已在应用套件中推出超过 1000 个 AI Agent。

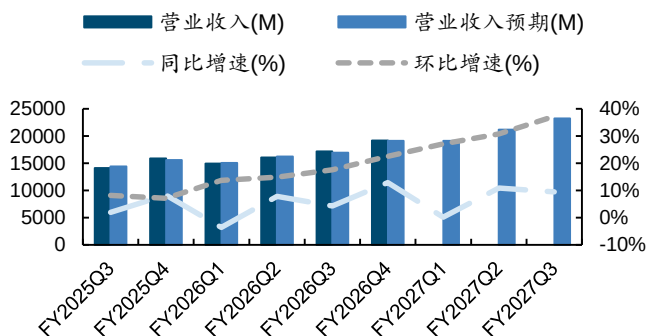
Oracle 短期毛利率承压, 主要因为数据中心建设和基础设施收入占比提升。全年毛利率下降约 5 个百分点, 2027 财年毛利率预计继续下降, 但管理层强调这更多是数据中心项目爬坡至完整收入贡献前的时间差与收入结构变化, 而非基础设施业务长期盈利能力恶化。

展望 2027 财年, Oracle 预计总收入按固定汇率增长 34% 以上, 超过此前长期收入 CAGR 目

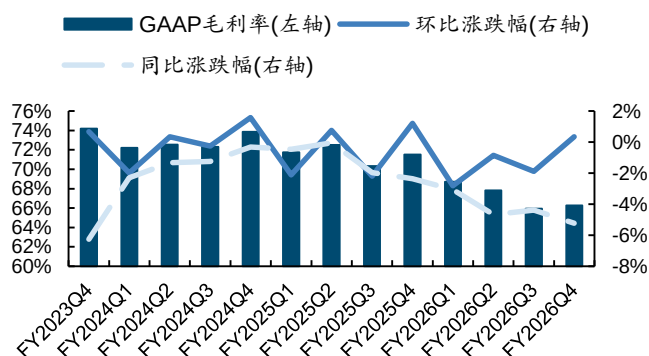


标。非 GAAP EPS 预计为 8.05 美元，按固定汇率增长 18%。2027 财年资本支出净现金流出预计约 700 亿美元，另有 200 亿至 250 亿美元客户预付款或时间性影响，因此实际的资本开支可能会更高。公司预计下一季度 (FY2027Q1) 收入同比增长 27%至 29%，云收入增长 58%至 64%，非 GAAP EPS 为 1.72-1.76 美元，并预计下半年随着更多数据中心容量上线，收入和利润增长将进一步加速。

图表3: Oracle 收入增速 (百万美元)



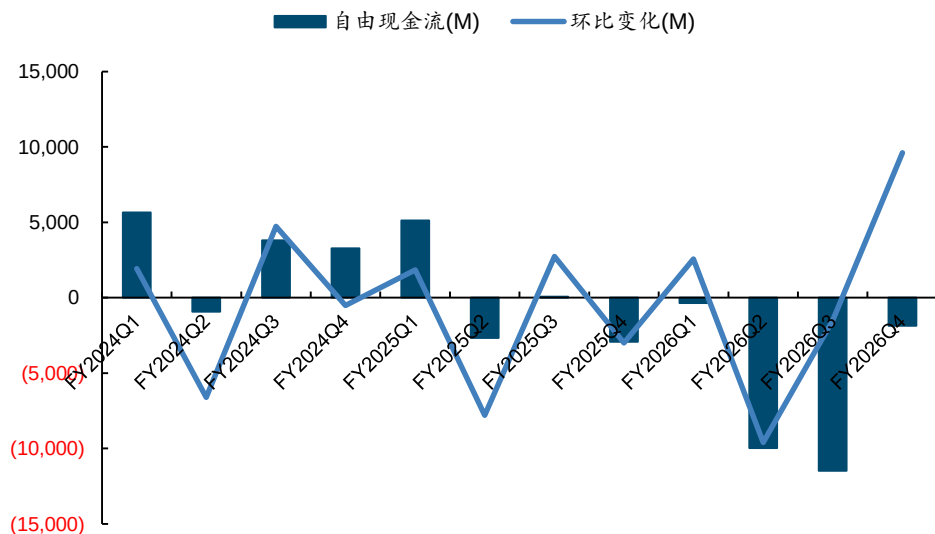
图表4: Oracle 毛利率



来源：路透、国金证券研究所

来源：路透、国金证券研究所

图表5: Oracle 自由现金流 (百万美元)



来源：路透、国金证券研究所

越来越多的人关注 Oracle 从软件公司转型至 AI 平台公司的过程中的盈利能力。目前 Oracle 除了以往的数据库以及 SaaS 业务等还参与到了数据中心建设。相比传统软件业务，数据中心基建项目前期投入较大、毛利率较低，业务的转型短期内会对 Oracle 经营产生影响。但是未来大的云厂商真正需求的并不只是 GPU，而是完整 AI 基础设施运营能力，包括电力、网络、存储、安全、身份管理以及超大规模 GPU 集群维护能力。我们认为，随着 Abilene、Shackelford、Dona Ana County 等多个 GW 级园区逐步上线，Oracle 收入与利润都将有所改善。

不过，未来几个季度市场仍需要重点观察 1) RPO 与 OCI 收入增速，这是 Oracle 未来增长最核心的领先指标，尤其需要观察未来 12 个月确认收入比例是否继续提升，以及 AI 基础设施收入能否维持高增长。2) 数据中心交付节奏，包括 Abilene 剩余容量上线进度、新园区是否延期以及 GPU 利用率是否维持高位。3) 基础设施利润率改善情况，目前毛利率受到数据中心建设与折旧压力影响。4) 资本开支与融资结构，数据中心建设需要大量资金，自由现金流与公司债利率影响到 Oracle 资金流动与融资难度。



风险提示

1. 芯片制程发展与良率不及预期：半导体工艺的发展面临诸多挑战，主要包括技术瓶颈、良率提升难度、研发成本高企以及供应链不确定性等问题。随着工艺节点微缩变得愈发复杂，先进制程的实现难度和成本不断攀升，可能导致量产延迟，甚至影响产品性能和成本控制。此外，地缘政治风险和出口管制可能扰乱供应链，进一步拖累产能扩张。
2. 中美科技领域政策恶化：中美在 AI 领域竞争激烈，美国限制先进芯片和半导体对中国的出口，随着竞争的加剧，未来可能会推出更严格的限制政策，限制国内 AI 模型的发展。
3. 智能手机销量不及预期：智能手机销量与产品本身质量关系紧密，若产品本身有缺陷则智能手机销量可能受到影响。同时宏观经济变化也有可能导致消费者消费意愿发生变化从而影响智能手机销量。



特别声明：

国金证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告版权归“国金证券股份有限公司”（以下简称“国金证券”）所有，未经事先书面授权，任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、转发、转载、引用、修改、仿制、刊发，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整，亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用，在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险，可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供多种金融服务。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密，只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级（含 C3 级）的投资者使用；本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具，本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告，则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议，国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有，保留一切权利。

上海	北京	深圳
电话：021-80234211	电话：010-85950438	电话：0755-86695353
邮箱：researchsh@gjzq.com.cn	邮箱：researchbj@gjzq.com.cn	邮箱：researchsz@gjzq.com.cn
邮编：201204	邮编：100005	邮编：518000
地址：上海浦东新区芳甸路 1088 号 紫竹国际大厦 5 楼	地址：北京市东城区建国内大街 26 号 新闻大厦 8 层南侧	地址：深圳市福田区金田路 2028 号皇岗商务中心 18 楼 1806



【小程序】
国金证券研究服务



【公众号】
国金证券研究